GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 30, 2002, 09:49:45 ; Search time 2325.7 Seconds
(without alignments)
992.189 Million cell updates/sec

Title: US-09-303-518D-131
Perfect score: 1344
Sequence: 1 atgattaaaatcaaaaaagg.....ccattgagaaggaaggctga 1344

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1344	100.0	1344	20	AAZ12028	Neisseria gonorrho
	2	1212.8	90.2	1344	20	AAZ12026	Neisseria meningit
c	3	1212.8	90.2	44608	21	AAA81495	N. meningitidis pa
c	4	1212.8	90.2	349980	21	AAF21607	Neisseria meningit
	5	1165.4	86.7	1344	20	AAZ12027	Neisseria meningit
	6	545.8	40.6	1830121	17	AAT42063	Haemophilus influe
	7	433.6	32.3	474	20	AAZ12025	Neisseria meningit
	8	433.6	32.3	474	21	AAA81335	N. meningitidis Me
c	9	363	27.0	363	21	AAZ54033	Neisseria gonorrhe
c	10	348.8	26.0	96109	22	AAF28548	Genomic fragment #
c	11	324.6	24.2	363	21	AAZ54034	Neisseria meningit
c	12	300.6	22.4	363	21	AAZ54035	Neisseria meningit
	13	126.4	9.4	1353	20	AAX91657	Porphorymonas ging
	14	126.4	9.4	1362	20	AAX91536	Porphorymonas ging
	15	38.8	2.9	125401	22	AAD17186	Streptomyces nours
	16	38.2	2.8	38734	20	AAZ32020	Human METH1 relate
	17	38.2	2.8	38734	22	AAC90077	AL021529 cDNA clon
	18	35.8	2.7	5036	22	AAD13450	Mycobacterium tube
	19	35.8	2.7	4403765	22	AAI99683	Mycobacterium tube
	20	35.8	2.7	4411529	22	AAI99682	Mycobacterium tube
c	21	35.6	2.6	5059	20	AAX84332	Stealth virus nucl
c	22	35.6	2.6	4403765	22	AAI99683	Mycobacterium tube
c	23	35.6	2.6	4411529	22	AAI99682	Mycobacterium tube
	24	35.4	2.6	296	23	AAS57519	cDNA #195 encoding
	25	35.4	2.6	10732	21	AAA10594	Gene encoding a su
	26	35.2	2.6	1938	23	AAS66442	DNA encoding novel
	27	35	2.6	323	23	AAS86095	DNA encoding novel
c	28	35	2.6	697	22	AAF25080	Nucleotide sequenc
	29	35	2.6	1139	21	AAF21735	Human breast and o
	30	35	2.6	1689	23	AAS86098	DNA encoding novel
	31	34.8	2.6	1325	22	AAD08837	Human G-protein co
c	32	34.8	2.6	1743	23	AAS85732	DNA encoding novel
c	33	34.8	2.6	1875	23	AAS85741	DNA encoding novel
	34	34.8	2.6	2196	23	AAS82534	DNA encoding novel
c	35	34.8	2.6	2196	23	AAS85733	DNA encoding novel
	36	34.8	2.6	2196	23	AAS89351	DNA encoding novel
	37	34.8	2.6	2909	23	AAS73293	DNA encoding novel
	38	34.8	2.6	3810	23	AAS82533	DNA encoding novel
	39	34.8	2.6	3810	23	AAS89350	Drosophila melanog
c	40	34.4	2.6	783	23	ABL02343	Drosophila melanog
c	41	34.4	2.6	2783	23	ABL02342	Propionibacterium
c	42	34.2	2.5	25365	23	AAS59558	Pseudomonas aerugi
	43	34	2.5	1191	23	AAS54057	DNA encoding novel
c	44	33.8	2.5	2367	23	AAS73398	DNA encoding novel
c	45	33.8	2.5	2367	23	AAS73427	DNA encoding novel
	46	33.8	2.5	2367	23	AAS88512	DNA encoding novel
c	47	33.8	2.5	2367	23	AAS93229	Drosophila melanog
c	48	33.6	2.5	1098	23	ABL02111	Human Zcytor14 var
c	49	33.6	2.5	1725	22	AAC85030	Human cytokine rec
c	50	33.6	2.5	2076	22	AAC85028	Human DCRS7 revers
c	51	33.6	2.5	2109	24	AAS18131	Drosophila melanog
	52	33.6	2.5	3475	23	ABL02112	Drosophila melanog
	53	33.6	2.5	4086	23	ABL02110	Enterococcus faeca
	54	33.6	2.5	7215	20	AAX13039	C glutamicum codin
c	55	33.4	2.5	1359	22	AAH65742	C glutamicum codin
	56	33.4	2.5	1773	22	AAH68405	Corynebacterium gl
	57	33.4	2.5	1896	22	AAF67754	C glutamicum codin
c	58	33.4	2.5	349980	22	AAH68526	Pseudomonas aerugi
	59	33.2	2.5	1488	23	AAS54225	DNA encoding novel
c	60	33.2	2.5	3747	23	AAS73044	Yeast AOD9604-asso
	61	33.2	2.5	4590	22	AAH24065	DNA encoding novel
c	62	33.2	2.5	4901	23	AAS92853	DNA encoding novel
c	63	33.2	2.5	6813	23	AAS85740	Human immune/haema
	64	33.2	2.5	23394	22	AAK77173	Nucleotide sequenc
c	65	33.2	2.5	77536	21	AAA14651	Mechanical stress
	66	33	2.5	462	21	AAZ36313	Trichoderma reesei
c	67	33	2.5	545	21	AAF15192	Pseudomonas aerugi
	68	33	2.5	783	23	AAS54375	Beta-amylase S291A
	69	33	2.5	1596	17	AAT03695	Plasmid pBETA92, e
	70	33	2.5	6312	17	AAT03696	Leishmania sp. ant
	71	32.8	2.4	1618	16	AAT05503	DNA encoding LbeIF
	72	32.8	2.4	1618	18	AAT80384	Leishmania brazili
	73	32.8	2.4	1618	18	AAT62139	Leishmania antigen
	74	32.8	2.4	1618	19	AAV47559	L. braziliensis EI
	75	32.8	2.4	1618	20	AAX02202	Leishmania antigen
	76	32.8	2.4	1618	24	AAS96023	N. meningitidis pa
	77	32.8	2.4	69936	21	AAA81479	Human protein HP03
	78	32.6	2.4	1116	22	AAF28683	Human G protein-co
	79	32.6	2.4	1116	24	ABA02176	DNA encoding a P2Y
	80	32.6	2.4	1119	21	AAA64367	Human orphan G pro
	81	32.6	2.4	1119	21	AAD01119	Human G protein co
	82	32.6	2.4	1119	21	AAA46018	

	83	32.6	2.4	1119	22	AAH49504	Human GTP-binding
	84	32.6	2.4	1119	22	AAD02585	Human G-protein co
	85	32.6	2.4	1119	22	AAF86237	Human G-protein co
	86	32.6	2.4	1119	24	AAS98045	Human DNA for pote
c	87	32.6	2.4	1237	24	AAS98085	Human DNA for pote
	88	32.6	2.4	1560	24	AAS19414	Human cDNA encodin
	89	32.6	2.4	1720	22	AAF28693	Human protein HP03
	90	32.6	2.4	2444	24	AAD26369	Human G-protein co
	91	32.6	2.4	2480	22	AAD06509	Human CON217 G pro
	92	32.6	2.4	2559	21	AAZ95039	Human G protein co
	93	32.6	2.4	3180	22	AAF25830	Human P2YL1 DNA.
	94	32.6	2.4	15783	22	AAS39803	Genomic sequence #
	95	32.6	2.4	15783	22	AAK90159	Human digestive sy
c	96	32.4	2.4	3412	20	AAX25772	S.erythraea erythr
	97	32.4	2.4	3756	18	AAT72684	Sugar biosynthesis
	98	32.4	2.4	3993	23	AAS88620	DNA encoding novel
c	99	32.4	2.4	5559	15	AAQ55260	Restriction fragme
c	100	32.4	2.4	5559	18	AAT99212	5.6 kb E11 Pseudom

ALIGNMENTS

RESULT 1
AAZ12028
ID AAZ12028 standard; DNA; 1344 BP.
XX
AC AAZ12028;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria gonorrhoeae ORF22 polynucleotide sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria gonorrhoeae.
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38565.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 9; Page 125; 524pp; English.
XX
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX

SQ Sequence 1344 BP; 334 A; 363 C; 362 G; 285 T; 0 other;

Query Match 100.0%; Score 1344; DB 20; Length 1344;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60
Db       1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60

Qy      61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120
Db      61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120

Qy     121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180
Db     121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy     181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240
Db     181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240

Qy     241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
Db     241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300

Qy     301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360
Db     301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360

Qy     361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420
Db     361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420

Qy     421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
Db     421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480

Qy     481 gccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggta 540
Db     481 gccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggta 540

Qy     541 ttgagccgcctgaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccg 600
Db     541 ttgagccgcctgaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccg 600

Qy     601 tctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttg 660
Db     601 tctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttg 660

Qy     661 agtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggaccatc 720
Db     661 agtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggaccatc 720

Qy     721 aattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaataccgag 780
Db     721 aattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaataccgag 780

Qy     781 cgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840
Db     781 cgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840

Qy     841 ggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtgatt 900
Db     841 ggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtgatt 900

Qy     901 tccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgctac 960
Db     901 tccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgctac 960

Qy     961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
```

```
Db    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020

Qy   1021 ccgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaa 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ccgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaa 1080

Qy   1081 ctcttcaagttcacgacagccgtcaacggcggcgaccgcgccatggtaccgatcggcact 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ctcttcaagttcacgacagccgtcaacggcggcgaccgcgccatggtaccgatcggcact 1140

Qy   1141 tatgagcgcgtaatgccgttggacatcctgcctaccttgcttttgcgcgatttaatcgtc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 tatgagcgcgtaatgccgttggacatcctgcctaccttgcttttgcgcgatttaatcgtc 1200

Qy   1201 ggcgataccgacagcgcgcaggctttgggttgcttggaattggacgaagaagacctcgct 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ggcgataccgacagcgcgcaggctttgggttgcttggaattggacgaagaagacctcgct 1260

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320

Qy   1321 gaaaccattgagaaggaaggctga 1344
          |||||||||||||||||||||||
Db   1321 gaaaccattgagaaggaaggctga 1344
```

RESULT 2
AAZ12026
ID AAZ12026 standard; DNA; 1344 BP.
XX
AC AAZ12026;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria meningitidis complete ORF22 sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria meningitidis.
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38561.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 9; Page 122-123; 524pp; English.
XX
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX
SQ Sequence 1344 BP; 329 A; 366 C; 365 G; 284 T; 0 other;

Query Match 90.2%; Score 1212.8; DB 20; Length 1344;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1262; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```
Qy      1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||| | ||
Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60

Qy     61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120
          || |||||||||||||||||||||||||||||||||||||||||||| ||| ||||||
Db     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180
          |||||||||||| |||||||||| || ||||||||||||||||||||||||||||||||
Db    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240
          |||||||||||||||| || |||||||||||||||||||||||||||||| ||||||
Db    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300

Qy    301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360
          ||||| |||||||||| |||||||||||||||||| || ||| ||||| ||||||||||
Db    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360

Qy    361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420
          ||||||||| ||||| || || |||||||||| |||||||||||||||||||||||| |||
Db    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420

Qy    421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480

Qy    481 gccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggta 540
          |||||||||||||||| |||||||||||||||| || ||||||||||||||||||||||
Db    481 gccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggta 540

Qy    541 ttgagccgcctgaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccg 600
          |||||||| |||||||||| |||||||||| ||||| ||||| |||||||||||||||||
Db    541 ttgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccg 600

Qy    601 tctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttg 660
          |||||||||||||||| |||||||||||||||| |||||||||||||||||||||| |||
Db    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttg 660

Qy    661 agtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggaccatc 720
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    661 agtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggaccatc 720

Qy    721 aattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaataccgag 780
          |||||||||| || ||| | || || |||||||| | ||||||||||||||| ||||||
Db    721 aattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacaccgag 780

Qy    781 cgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840
          |||||| ||||| | || ||    ||||||||||||||||||||||||||||||||||
Db    781 cgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840

Qy    841 ggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtgatt 900
          |||||||| || || ||| |||| || ||||||||||||||| | ||||||||||||||
Db    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacacagacaaccgcgtgatt 900

Qy    901 tccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgctac 960
          |||||||||||||||||||| |||||| |||||||||||| ||||||||||||||||||
```

```
Db    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020

Qy   1021 ccgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaa 1080
          ||||||||||||||||||||||||||| || || ||||||||||||| |||||||||
Db   1021 ccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcctgaaaaacaaa 1080

Qy   1081 ctcttcaagttcacgacagccgtcaacggcggcgaccgcgccatggtaccgatcggcact 1140
          |||||||||||||  ||||||||||||||||||||||||||||||| ||||| || |||
Db   1081 ctcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtgccgattggtact 1140

Qy   1141 tatgagcgcgtaatgccgttggacatcctgcctaccttgcttttgcgcgatttaatcgtc 1200
          || |||||||| ||||| ||||| |||||||| ||| ||||||||||||||||||||||
Db   1141 tacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgcgatttaatcgtc 1200

Qy   1201 ggcgataccgacagcgcgcaggctttgggttgcttggaattggacgaagaagacctcgct 1260
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   1201 ggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaagaagacctcgct 1260

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320

Qy   1321 gaaaccattgagaaggaaggctga 1344
          |||||||||||||||||||||||
Db   1321 gaaaccattgagaaggaaggctga 1344
```

```
RESULT   3
AAA81495/c
ID   AAA81495 standard; DNA; 44608 BP.
XX
AC   AAA81495;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_42 SEQ ID NO:42.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1283-1296; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 44608 BP; 10938 A; 10835 C; 11999 G; 10834 T; 2 other;
```

```
  Query Match             90.2%;  Score 1212.8;  DB 21;  Length 44608;
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches 1262;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy      1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||| | ||
Db  25944 ATGATTAAAATCAAAAAAGGTCTAAACCTGCCCATCGCGGGCAGACCGGAGCAAGCCGTT 25885

Qy     61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120
          || |||||||||||||||||||||||||||||||||||||||||||||||| ||| |||||||
Db  25884 TACGACGGCCCGGCCATTACCGAAGTCGCGTTGCTTGGCGAAGAATATGCCGGTATGCGC 25825

Qy    121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180
          |||||||||||| |||||||||| || ||||||||||||||||||||||||||||||||||
Db  25824 CCCTCGATGAAAGTCAAGGAAGGCGATGCCGTCAAAAAAGGCCAAGTGCTGTTTGAAGAC 25765

Qy    181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240
          |||||||||||||||| || ||||||||||||||||||||||||||||||||| ||||||
Db  25764 AAAAAGAATCCGGGCGTGGTGTTTACTGCGCCGGCTTCAGGCAAAATCGCCGCGATTCAC 25705

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  25704 CGTGGCGAAAAGCGCGTACTTCAGTCAGTCGTGATTGCCGTTGAAGGCAACGACGAAATC 25645

Qy    301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360
          ||||| |||||||||| ||||||||||||||||| || ||| ||||| ||||||||||
Db  25644 GAGTTTGAACGCTACGCACCTGAAGCGCTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGC 25585

Qy    361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420
          |||||||| ||||| || || |||||||||||| |||||||||||||||||||||||| |||
Db  25584 AACCTGATCCAATCCGGTTTGTGGACTGCGCTGCGCACCCGTCCGTTCAGCAAAATTCCT 25525

Qy    421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  25524 GCCGTCGATGCCGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCT 25465

Qy    481 gccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggta 540
          |||||||||||||||| |||||||||||||||| || ||||||||||||||||||||
Db  25464 GCCGACCCTACGGTCATTATCAAAGAAGCCGCCGAGGATTTCAAACGCGGCCTGTTGGTA 25405

Qy    541 ttgagccgcctgaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccg 600
          ||||||||  |||||||||| |||||||||| ||||| ||||| ||||||||||||||||
Db  25404 TTGAGCCGTTTGACCGAACGCAAAATCCATGTTTGTAAGGCAGCTGGCGCAGACGTGCCG 25345

Qy    601 tctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttg 660
          |||||||||||||||| |||||||||||||||||| |||||||||||||||||||| |||
Db  25344 TCTGAAAATGCTGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCTGCCGGTTTG 25285
```

```
Qy    661 agtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggaccatc 720
          ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  25284 AGTGGCACGCACATTCATTTCATCGAGCCGGTCGGCGCGAATAAAACCGTGTGGACCATC 25225

Qy    721 aattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaataccgag 780
          |||||||||| || ||| | || || |||||||| | ||||||||||||||| ||||||
Db  25224 AATTATCAAGATGTAATTACCATTGGCCGTTTGTTTGCAACAGGCCGTCTGAACACCGAG 25165

Qy    781 cgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840
          |||||| ||||| | || ||    ||||||||||||||||||||||||||||||||||||
Db  25164 CGCGTGATTGCCCTAGGTGGTTCTCAAGTCAACAAACCGCGCCTCTTGCGTACCGTTTTG 25105

Qy    841 ggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtgatt 900
          |||||||| || || ||| |||| || ||||||||||||||| | ||||||||||||||
Db  25104 GGTGCGAAAGTATCGCAAATTACTGCGGGCGAATTGGTTGACACAGACAACCGCGTGATT 25045

Qy    901 tccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgctac 960
          |||||||||||||||||||| |||||| |||||||||||||| ||||||||||||||||||
Db  25044 TCCGGTTCGGTATTGAACGGCGCGATTACACAAGGCGCGCACGATTATTTGGGACGCTAC 24985

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  24984 CACAATCAGATTTCCGTTATCGAAGAAGGCCGCAGCAAAGAGCTGTTCGGCTGGGTTGCG 24925

Qy   1021 ccgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaa 1080
          |||||||||||||||||||||||||||| || || |||||||||||||| |||||||||
Db  24924 CCGCAGCCGGACAAATACTCCATCACGCGTACAACCCTCGGCCATTTCCTGAAAAACAAA 24865

Qy   1081 ctcttcaagttcacgacagccgtcaacggcggcgaccgcgccatggtaccgatcggcact 1140
          |||||||||||||  |||||||||||||||||||||||||||||| ||||| || |||
Db  24864 CTCTTCAAGTTCAACACAGCCGTCAACGGCGGCGACCGCGCCATGGTGCCGATTGGTACT 24805

Qy   1141 tatgagcgcgtaatgccgttggacatcctgcctaccttgcttttgcgcgatttaatcgtc 1200
          || |||||||| ||||| ||||| |||||||| ||| ||||||||||||||||||||||
Db  24804 TACGAGCGCGTGATGCCCTTGGATATCCTGCCCACCCTGCTTTTGCGCGATTTAATCGTC 24745

Qy   1201 ggcgataccgacagcgcgcaggctttgggttgcttggaattggacgaagaagacctcgct 1260
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  24744 GGCGATACCGACAGCGCGCAGGCATTGGGTTGCTTGGAATTGGACGAAGAAGACCTCGCT 24685

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  24684 TTGTGCAGCTTCGTCTGCCCGGGCAAATACGAATACGGCCCGCTGTTGCGCAAAGTGCTG 24625

Qy   1321 gaaaccattgagaaggaaggctga 1344
          |||||||||||||||||||||||
Db  24624 GAAACCATTGAGAAGGAAGGCTGA 24601
```

```
RESULT   4
AAF21607/c
ID   AAF21607 standard; DNA; 349980 BP.
XX
AC   AAF21607;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Neisseria meningitidis B nucleotide sequence SEQ ID NO:108.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW   diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW   ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200066791-A1.
XX
PD   09-NOV-2000.
XX
PF   08-MAR-2000; 2000WO-US05928.
XX
PR   30-APR-1999;   99US-0132068.
PR   08-OCT-1999;   99WO-US23573.
PR   28-FEB-2000; 2000GB-0004695.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Pizza M,  Hickey E,  Peterson J,  Tettelin H,  Venter JC,  Masignani V;
PI   Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V,  Rappuoli R;
PI   Frazer CM,  Grandi G;
XX
DR   WPI; 2000-647603/62.
XX
PT   Neisseria meningitidis B full length genome sequence and open reading
PT   frames are used to detect, treat and prevent Neisserial infections -
XX
PS   Claim 7; Appendix A; 692pp; English.
XX
CC   The present invention describes the full length genome of
CC   Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC   to AAF21613 represent fragments of the NMB genomic sequence, as the
CC   sequence was too long to go in a record on its own it was split into 8
CC   sequences which overlap each other at the beginning and end of each
CC   sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC   the beginning  of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC   the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC   Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC   AAF21606 represent PCR primers which are used in the exemplification of
CC   the present invention. The NMB genome and fragments from it have
CC   antibacterial activity, and can be used in vaccines and gene therapy.
CC   Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC   proteins can be used in compositions for treating or preventing infection
CC   due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC   presence of Neisserial bacteria or of antibodies raised to Neisserial
CC   bacteria. Computers, computer memory, computer storage medium or computer
CC   databases can be used in a search to identify open reading frames (ORFs)
CC   or coding sequences within the NMB genome. The DNA sequences provide
CC   further opportunities to find antigenic or immunogenic proteins which are
CC   more effective in vaccines than the outer membrane proteins currently
CC   used.
XX
SQ   Sequence 349980 BP; 84410 A; 84863 C; 94187 G; 86520 T; 0 other;
```

```
  Query Match             90.2%;  Score 1212.8;  DB 21;  Length 349980;
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches 1262;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy       1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||| | ||
Db  297909 ATGATTAAAATCAAAAAAGGTCTAAACCTGCCCATCGCGGGCAGACCGGAGCAAGCCGTT 297850

Qy      61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120
           || |||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||||
Db  297849 TACGACGGCCCGGCCATTACCGAAGTCGCGTTGCTTGGCGAAGAATATGCCGGTATGCGC 297790

Qy     121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180
           |||||||||||| |||||||||| || ||||||||||||||||||||||||||||||||||
Db  297789 CCCTCGATGAAAGTCAAGGAAGGCGATGCCGTCAAAAAAGGCCAAGTGCTGTTTGAAGAC 297730

Qy     181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240
           |||||||||||||||| || |||||||||||||||||||||||||||||||||| ||||||
Db  297729 AAAAAGAATCCGGGCGTGGTGTTTACTGCGCCGGCTTCAGGCAAAATCGCCGCGATTCAC 297670

Qy     241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  297669 CGTGGCGAAAAGCGCGTACTTCAGTCAGTCGTGATTGCCGTTGAAGGCAACGACGAAATC 297610

Qy     301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360
           ||||| |||||||||| |||||||||||||||||| || ||| ||||| ||||||||||
Db  297609 GAGTTTGAACGCTACGCACCTGAAGCGCTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGC 297550

Qy     361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420
           |||||||| ||||| || || |||||||||||| |||||||||||||||||||||||| |||
```

```
Db 297549 AACCTGATCCAATCCGGTTTGTGGACTGCGCTGCGCACCCGTCCGTTCAGCAAAATTCCT 297490

Qy     421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297489 GCCGTCGATGCCGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCT 297430

Qy     481 gccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggta 540
           |||||||||||||||| |||||||||||||||| || ||||||||||||||||||||||||
Db 297429 GCCGACCCTACGGTCATTATCAAAGAAGCCGCCGAGGATTTCAAACGCGGCCTGTTGGTA 297370

Qy     541 ttgagccgcctgaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccg 600
           |||||||| |||||||||| |||||||||| ||||| ||||| ||||||||||||||||
Db 297369 TTGAGCCGTTTGACCGAACGCAAAATCCATGTTTGTAAGGCAGCTGGCGCAGACGTGCCG 297310

Qy     601 tctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttg 660
           |||||||||||||||| |||||||||||||||| |||||||||||||||||||||||| |||
Db 297309 TCTGAAAATGCTGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCTGCCGGTTTG 297250

Qy     661 agtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggaccatc 720
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 297249 AGTGGCACGCACATTCATTTCATCGAGCCGGTCGGCGCGAATAAAACCGTGTGGACCATC 297190

Qy     721 aattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaatacccgag 780
           |||||||||||| || ||| | || || |||||||| | |||||||||||||||| ||||||
Db 297189 AATTATCAAGATGTAATTACCATTGGCCGTTTGTTTGCAACAGGCCGTCTGAACACCGAG 297130

Qy     781 cgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840
           |||||| ||||| | || ||    ||||||||||||||||||||||||||||||||||||||
Db 297129 CGCGTGATTGCCCTAGGTGGTTCTCAAGTCAACAAACCGCGCCTCTTGCGTACCGTTTTG 297070

Qy     841 ggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtgatt 900
           |||||||| || || ||| |||| || |||||||||||||| | ||||||||||||||||
Db 297069 GGTGCGAAAGTATCGCAAATTACTGCGGGCGAATTGGTTGACACAGACAACCGCGTGATT 297010

Qy     901 tccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgctac 960
           |||||||||||||||||||| |||||| |||||||||||| ||||||||||||||||||||
Db 297009 TCCGGTTCGGTATTGAACGGCGCGATTACACAAGGCGCGCACGATTATTTGGGACGCTAC 296950

Qy     961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296949 CACAATCAGATTTCCGTTATCGAAGAAGGCCGCAGCAAAGAGCTGTTCGGCTGGGTTGCG 296890

Qy    1021 ccgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaa 1080
           |||||||||||||||||||||||||||||| || || |||||||||||||| ||||||||||
Db 296889 CCGCAGCCGGACAAATACTCCATCACGCGTACAACCCTCGGCCATTTCCTGAAAAACAAA 296830

Qy    1081 ctcttcaagttcacgacagccgtcaacggcggcgaccgcgccatggtaccgatcggcact 1140
           |||||||||||||  ||||||||||||||||||||||||||||||||||||| || |||
Db 296829 CTCTTCAAGTTCAACACAGCCGTCAACGGCGGCGACCGCGCCATGGTGCCGATTGGTACT 296770

Qy    1141 tatgagcgcgtaatgccgttggacatcctgcctaccttgcttttgcgcgatttaatcgtc 1200
           || |||||||| ||||| ||||| |||||||| ||| ||||||||||||||||||||||||
Db 296769 TACGAGCGCGTGATGCCCTTGGATATCCTGCCCACCCTGCTTTTGCGCGATTTAATCGTC 296710

Qy    1201 ggcgataccgacagcgcgcaggctttgggttgcttggaattggacgaagaagacctcgct 1260
           ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 296709 GGCGATACCGACAGCGCGCAGGCATTGGGTTGCTTGGAATTGGACGAAGAAGACCTCGCT 296650

Qy    1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296649 TTGTGCAGCTTCGTCTGCCCGGGCAAATACGAATACGGCCCGCTGTTGCGCAAAGTGCTG 296590

Qy    1321 gaaaccattgagaaggaaggctga 1344
           |||||||||||||||||||||||||
Db 296589 GAAACCATTGAGAAGGAAGGCTGA 296566


RESULT   5
AAZ12027
ID   AAZ12027 standard; DNA; 1344 BP.
XX
AC   AAZ12027;
XX
DT   08-OCT-1999  (first entry)
XX
DE   Neisseria meningitidis strain A complete ORF22 sequence.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9924578-A2.
XX
PD   20-MAY-1999.
XX
PF   09-OCT-1998;   98WO-IB01665.
XX
PR   01-SEP-1998;   98GB-0019016.
PR   06-NOV-1997;   97GB-0023516.
PR   14-NOV-1997;   97GB-0024190.
PR   18-NOV-1997;   97GB-0024386.
PR   27-NOV-1997;   97GB-0025158.
PR   10-DEC-1997;   97GB-0026147.
PR   14-JAN-1998;   98GB-0000759.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Grandi G,  Masignani V,  Pizza M,  Rappuoli R,  Scarlato V;
XX
DR   WPI; 1999-327407/27.
DR   P-PSDB; AAY38562.
XX
PT   Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT   diagnosis, treatment and prevention of infection
XX
PS   Claim 9; Page 123; 524pp; English.
XX
CC   Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC   (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC   antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC   fragments, their nucleic acids and antibodies are used for diagnosis,
CC   prevention (as vaccines) or treatment of Neisseria infections,
CC   such as meningitis, septicaemia and gonorrhea. Both organisms
CC   are closely related. Fragments of the nucleic acids are useful
CC   as hybridisation probes and antisense reagents.
XX
SQ   Sequence 1344 BP; 323 A; 347 C; 368 G; 289 T; 17 other;


  Query Match             86.7%;  Score 1165.4;  DB 20;  Length 1344;
  Best Local Similarity   91.2%;  Pred. No. 0;
  Matches 1226;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy       1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60
           ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60

Qy      61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120
           |||||||| || | ||||||||||||||||||||||||||||||||| ||| ||||||
Db      61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy     121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180
           |||| ||||||| |||||||||| || ||||||||||||||||||||||||||||||||||
Db     121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy     181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240
           |||||| |||||||||| || ||||| |||| | ||||||||||||||||||| || ||
Db     181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240

Qy     241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
           || |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300

Qy     301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360
           ||||||||||||||||  || |||||| ||||||| || ||| ||||  || |  |  ||
```

```
Db    301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngc 360

Qy    361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420
          || |||||| ||||| || || |||||||||||| || | ||||||||||||||||||||||
Db    361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420

Qy    421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcg 480

Qy    481 gccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggta 540
          || |||||| ||| | |||||||||||||| ||| || |||| ||| | || ||||||
Db    481 gcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctggta 540

Qy    541 ttgagccgcctgaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccg 600
          |||||||| ||||||| |||||||||||||||||||| ||||| ||||||||||||||||
Db    541 ttgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccg 600

Qy    601 tctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttg 660
          |||||||||||||||| |||||||||||||||| |||||||||||||| ||||| |||
Db    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttg 660

Qy    661 agtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggaccatc 720
          |||||||||||||||||||||| ||||| ||||| || || |||||||| |||||||||
Db    661 agtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggaccatc 720

Qy    721 aattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaataccgag 780
          |||||||||| || ||||| |||||||||||||| | |||||||||||||| ||||||
Db    721 aattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacaccgag 780

Qy    781 cgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840
          |||||| |||| ||||| ||    |||||||||||||| ||||||||||||||||||||
Db    781 cgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctcttgcgtaccgttttg 840

Qy    841 ggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtgatt 900
          |||||||| || || ||| |||| || |||||||||||||||| ||||||||||||||||
Db    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacgcagacaaccgcgtgatt 900

Qy    901 tccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgctac 960
          |||||||||||||||||||| |||||| |||||||||||||| ||||||||||||||||||
Db    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020

Qy   1021 ccgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaa 1080
          |||||||||||||||||||||||||||| || || |||||||||||||| |||||||||
Db   1021 ccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcctgaaaaacaaa 1080

Qy   1081 ctcttcaagttcacgacagccgtcaacggcggcgaccgcgccatggtaccgatcggcact 1140
          |||||||||||||||||||||||||||| |||||||||||||||| ||||| || |||
Db   1081 ctcttcaagttcacgacagccgtcaacggtggcgaccgcgccatggtgccgattggtact 1140

Qy   1141 tatgagcgcgtaatgccgttggacatcctgcctaccttgcttttgcgcgatttaatcgtc 1200
          || |||||||||||||| | |||||||||||||||| ||||||||||||||||||||||
Db   1141 tacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgcgcgatttaatcgtc 1200

Qy   1201 ggcgataccgacagcgcgcaggctttgggttgcttggaattggacgaagaagacctcgct 1260
          |||||||||||||||||| || ||||||||||||||||||||||||||||||||||||||
Db   1201 ggcgataccgacagcgcgcaagcattgggttgcttggaattggacgaagaagacctcgct 1260

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320
          |||||||||||||||||||||||||||||||||||| |||||||||||||| || ||||||
Db   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgttgcgtaaggtgctg 1320

Qy   1321 gaaaccattgagaaggaaggctga 1344
          |||||| ||||||||||||||||
Db   1321 gaaaccnttgagaaggaaggctga 1344

RESULT 6
AAT42063
```

```
ID   AAT42063 standard; DNA; 1830121 BP.
XX
AC   AAT42063;
XX
DT   14-SEP-1999  (first entry)
XX
DE   Haemophilus influenzae complete genome sequence.
XX
KW   Genome; bacterium; Haemophilus influenzae; computer readable medium;
KW   expression modulating fragment; regulation; gene expression; vector;
KW   organism; open reading frame; ORF; ds.
XX
OS   Haemophilus influenzae.
XX
PN   WO9633276-A1.
XX
PD   24-OCT-1996.
XX
PF   22-APR-1996;   96WO-US05320.
XX
PR   07-JUN-1995;   95US-0487429.
PR   21-APR-1995;   95US-0426787.
PR   07-JUN-1995;   95US-0476102.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Adams MD,  Fleischmann RD,  Smith HO,  Venter JC,  White O;
XX
DR   WPI; 1996-485782/48.
XX
PT   Haemophilus influenzae Rd genome recorded on computer readable
PT   medium - useful for identifying commercially important nucleic acid
PT   fragments by homology searching
XX
PS   Claim 1; Page 77.2-77.1091; 1291pp; English.
XX
CC   This sequence represents the complete genome sequence of the bacterium
CC   Haemophilus influenzae strain Rd.  The invention relates to a computer
CC   readable medium (CRM) having recorded upon it the complete H.influenzae
CC   nucleotide sequence (I), a representative fragment of (I) or a nucleotide
CC   sequence at least 99% identical to (I).  By providing the full-length
CC   genomic sequence in a computer readable form, it is possible to identify
CC   commercially important nucleic acid fragments and expression modulating
CC   fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
CC   regulate the expression of a nucleic acid molecule.  Vectors and altered
CC   organisms comprising the predicted ORFS can be used to produce any of the
CC   polypeptide fragments of the H. influenzae Rd genome.
XX
SQ   Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;
```

```
  Query Match             40.6%;  Score 545.8;  DB 17;  Length 1830121;
  Best Local Similarity   63.8%;  Pred. No. 4.9e-155;
  Matches  860;  Conservative    0;  Mismatches  482;  Indels    5;  Gaps    2;

Qy       1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60
           ||||||| ||| || |||||| |  |||| || || ||||| | ||| |  ||||| ||
Db  179007 atgattacaattaagaaaggtttggatctcccaattgcgggaaaaccagcacaagtaatc 179066

Qy      61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120
             ||  ||||   ||  | |   |||| ||| | || || || || ||||| || ||||||
Db  179067 catagcggcaacgctgtgaatcaagttgcgattctaggtgaggagtatgtggggatgcgt 179126

Qy     121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180
           || || |||||  |    ||||| || |  || |||||||| ||||| || ||||||||||
Db  179127 ccttcaatgaaggtgcgcgaaggcgatgttgtgaaaaaaggtcaagtacttttgaagac 179186

Qy     181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240
           ||||| ||||| || ||| | ||||| || || ||    || |  ||| | || ||  |
Db  179187 aaaaaaaatcctggtgtaatttttacagcccctgcaagcggtaccatcactgcaatcaat 179246

Qy     241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
```

```
Db 179247 cgtggcgaaaaacgtgtattacaatctgttgttattaatgtagaaggtgatgaaaaaatc 179306

Qy    301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360
Db 179307 acttttgcaaaatatagcacagaacaattgaatacactttcttctgaacaagtaaaacaa 179366

Qy    361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420
Db 179367 aacttgatagaatctggcttatggacggcattacgtactcgtccatttagtaaagttcca 179426

Qy    421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
Db 179427 tccattgaaagtgaagcatcttctatttttgtaaatgcgatggataccaatccattagca 179486

Qy    481 gccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggta 540
Db 179487 gccgatccctctgtcgtattaaaagaatattcgcaagacttcactaacggtttaaccgta 179546

Qy    541 ttgagccgcctgaccgaacgtaa---aatccatgtgtgtaaagcagcaggcgcagacgtg 597
Db 179547 ttaagtcgtctattcccttcaaaaccattacacttatgtaaagcaggcgattctaatatt 179606

Qy    598 ccgtctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggc 657
Db 179607 ccaactgctgatcttgaaaatttacagattcatgattttactggtgttcatccagctggt 179666

Qy    658 ttgagtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggacc 717
Db 179667 cttgtagggacgcatattcactttattgatcctgtcggtattcaaaaaactgtatggcat 179726

Qy    718 atcaattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaatacc 777
Db 179727 atcaattatcaagatgtgatcgctgtaggtaaattatttacaacaggcgaactttattca 179786

Qy    778 gagcgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgtt 837
Db 179787 gaacgagtaatttctcttgcaggtcctcaagtgaaagaacctcgtttagtacgtacaaca 179846

Qy    838 ttgggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtg 897
Db 179847 atcggtgctaatctttctcaattaacccaaaatgaattaagtgcaggtaaaaatcgtgta 179906

Qy    898 atttccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgc 957
Db 179907 atttctggctcagttctttgtggtcaaatagcaaaagattcacatgattatttaggtcgc 179966

Qy    958 taccacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggtt 1017
Db 179967 tatgctttgcaagtatcagtaatcgctgaaggtaatgaaaaagaatttttggttggatt 180026

Qy   1018 gcgccgcagccggacaaatactccatcacgcgcaccactctcggccatttoctaaaaaac 1077
Db 180027 atgccacaagccaataaatattctgtgactcgtaccgtattagggc--cacttcagtaaa 180084

Qy   1078 aaactcttcaagttcacgacagccgtcaacggcggcgaccgcgccatggtaccgatcggc 1137
Db 180085 aaattgtttaattttactacttcagaaaatggtggcgaacgtgcaatggtgccaatcggt 180144

Qy   1138 acttatgagcgcgtaatgccgttggacatcctgcctaccttgcttttgcgcgatttaatc 1197
Db 180145 agctatgagcgcgtaatgccattggacattttacctacattattattacgtgatttaatc 180204

Qy   1198 gtcggcgataccgacagcgcgcaggctttgggttgcttggaattggacgaagaagacctc 1257
Db 180205 gtgggcgataccgatggtgctcaagagttaggttgtctggaattagacgaagaagacttg 180264

Qy   1258 gctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtg 1317
Db 180265 gcgttatgttctttcgtttgtccaggcaaatatgaatacggttcaatcttgcgtcaagtc 180324

Qy   1318 ctggaaaccattgagaaggaaggctga 1344
```

```
Db 180325 ttagataagattgagaaggaaggttaa 180351
```

RESULT 7
AAZ12025
ID AAZ12025 standard; DNA; 474 BP.
XX
AC AAZ12025;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria meningitidis partial ORF22 sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria meningitidis.
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38560.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 9; Page 122; 524pp; English.
XX
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX
SQ Sequence 474 BP; 122 A; 130 C; 130 G; 91 T; 1 other;

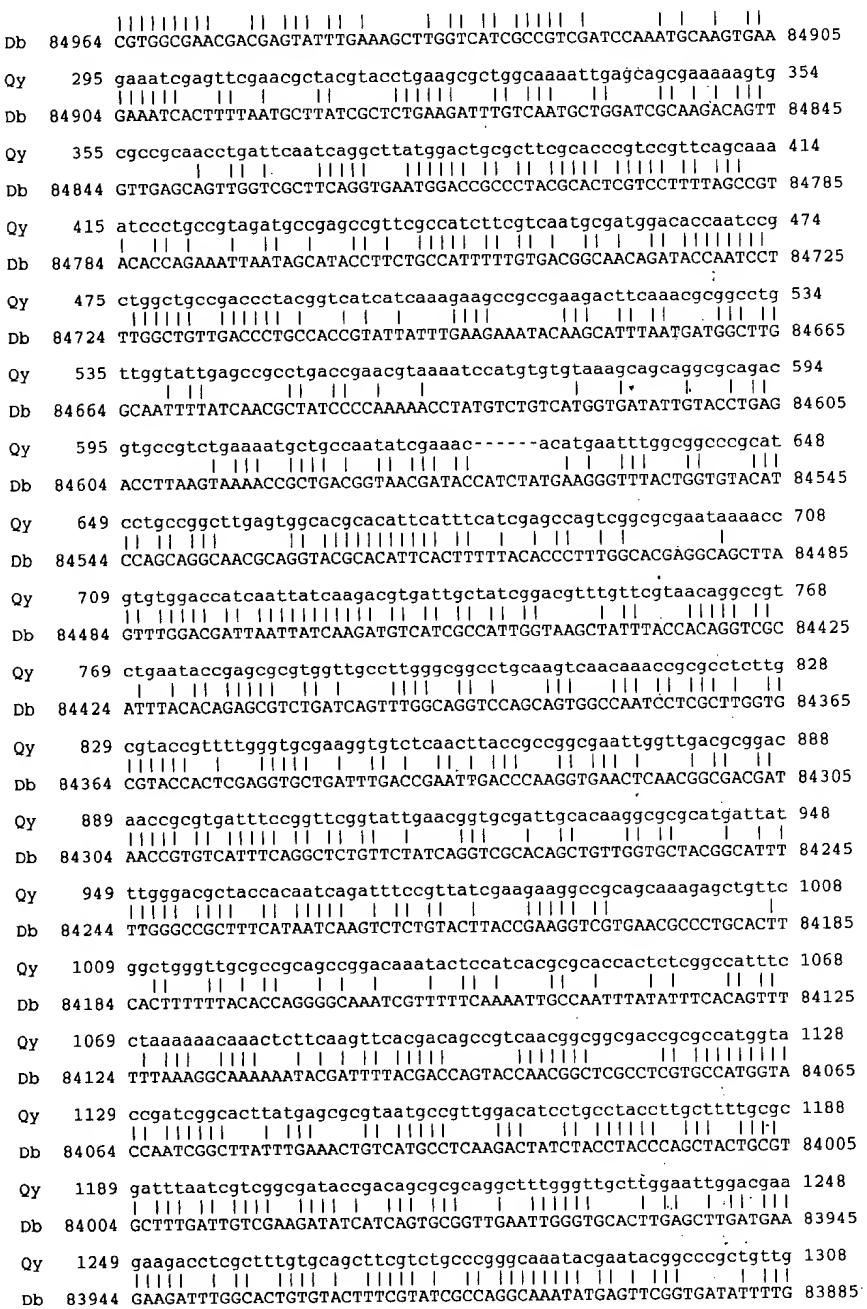
Query Match 32.3%; Score 433.6; DB 20; Length 474;
Best Local Similarity 94.5%; Pred. No. 2e-122;
Matches 448; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

```
Qy      1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60
Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60

Qy     61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120
Db     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180
Db    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240
```

```
                ||||||||||||||||| || |||||||||||||||||||||||||||||||||| ||||||
Db    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaargcaacgacgaaatc 300

Qy    301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360
          ||||| |||||||||| |||||||||||||||||| || ||| ||||| ||||||||||
Db    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360

Qy    361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420
          |||||||| ||||| || || |||||||||| ||||||||||||||||||||||| |||
Db    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420

Qy    421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474


RESULT   8
AAA81335
ID   AAA81335 standard; DNA; 474 BP.
XX
AC   AAA81335;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis MenB polynucleotide sequence ORF number 14.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Disclosure; Page 203; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 474 BP; 122 A; 130 C; 130 G; 91 T; 1 other;


  Query Match             32.3%;  Score 433.6;  DB 21;  Length 474;
  Best Local Similarity   94.5%;  Pred. No. 2e-122;
  Matches  448;  Conservative    1;  Mismatches   25;  Indels    0;  Gaps    0;

Qy      1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||| | ||
Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60

Qy     61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120
          || ||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||||
Db     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180
          |||||||||||| |||||||||| || ||||||||||||||||||||||||||||||||
Db    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240
          |||||||||||||||||| || |||||||||||||||||||||||||||||| ||||||
Db    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
          ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaargcaacgacgaaatc 300

Qy    301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360
          ||||| |||||||||| |||||||||||||||||| || ||| ||||| ||||||||||
Db    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360

Qy    361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420
          |||||||| ||||| || || |||||||||| ||||||||||||||||||||||| |||
Db    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420

Qy    421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474


RESULT   9
AAZ54033/c
ID   AAZ54033 standard; DNA; 363 BP.
XX
AC   AAZ54033;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria gonorrheae ORF 628 partial DNA sequence SEQ ID NO:2015.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria gonorrheae.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
```

```
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY75271.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics -
XX
PS   Claim 7; Page 1003; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 363 BP; 71 A; 92 C; 109 G; 91 T; 0 other;

  Query Match             27.0%;  Score 363;  DB 21;  Length 363;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-101;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    312 ctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgcaacctgattca 371
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 CTACGTACCTGAAGCGCTGGCAAAATTGAGCAGCGAAAAAGTGCGCCGCAACCTGATTCA 304

Qy    372 atcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccctgccgtagatgc 431
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 ATCAGGCTTATGGACTGCGCTTCGCACCCGTCCGTTCAGCAAAATCCCTGCCGTAGATGC 244

Qy    432 cgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggctgccgaccctac 491
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 CGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCTGCCGACCCTAC 184

Qy    492 ggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggtattgagccgcct 551
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 GGTCATCATCAAAGAAGCCGCCGAAGACTTCAAACGCGGCCTGTTGGTATTGAGCCGCCT 124

Qy    552 gaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccgtctgaaaatgc 611
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 GACCGAACGTAAAATCCATGTGTGTAAAGCAGCAGGCGCAGACGTGCCGTCTGAAAATGC 64

Qy    612 tgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttgagtggcacgca 671
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 TGCCAATATCGAAACACATGAATTTGGCGGCCCGCATCCTGCCGGCTTGAGTGGCACGCA 4

Qy    672 cat 674
          |||
Db      3 CAT 1
```

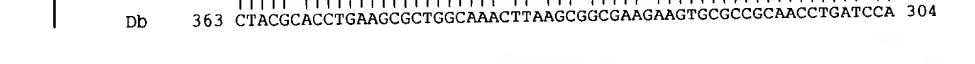
```
RESULT  10
AAF28548/c
ID   AAF28548 standard; DNA; 96109 BP.
XX
AC   AAF28548;
XX
DT   04-APR-2001  (first entry)
XX
DE   Genomic fragment #35.
XX
KW   Genomic library; bacteria; human upper airway; otitis media; sinusitis;
KW   bronchopulmonary; endocarditis; meningitis; ss.
XX
OS   Moraxella catarrhalis.
XX
PN   WO200078968-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16649.
XX
PR   18-JUN-1999;   99US-0140121.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lagace RE,  Patterson C,  Berg KL;
XX
DR   WPI; 2001-041427/05.
XX
PT   Genomic library for identifying diagnostic and therapeutic
PT   compositions, and for identifying virulence factors, regulatory
PT   elements and drug targets, comprises Moraxella catarrhalis nucleic
PT   acids -
XX
PS   Claim 1; Page 345-368; 545pp; English.
XX
CC   The present invention relates to a Moraxella catarrhalis genomic library
CC   comprising of a combination of 41 nucleic acid molecules (see
CC   AAF28514-AAF28554). The library has a number of uses described in the
CC   specification e.g. is useful for aidentifying diagnostic and therapeutic
CC   compositions. M. catarrhalis (Branhamella catarrhalis) is a large
CC   aerobic, gram-negative diplococcus, normally found among the bacterial
CC   flora of human upper airways. M. catarrhalis is known to cause acute,
CC   localised infections such as otitis media, sinusitis and bronchopulmonary
CC   infection and life-threatening, systemic diseases including endocarditis
CC   and meningitis.
XX
SQ   Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 other;

  Query Match             26.0%;  Score 348.8;  DB 22;  Length 96109;
  Best Local Similarity   55.2%;  Pred. No. 2.5e-95;
  Matches  749;  Conservative    0;  Mismatches  592;  Indels   15;  Gaps    3;

Qy      1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60
          |||||||  ||||||||||||  |  ||||||| ||| | |||   ||     |   |||
Db  85201 ATGATTACCATCAAAAAAGGCTTGGATCTGCCAATCACTGGCGAGCCCTCCAAGCAGATT 85142

Qy     61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120
            ||   | || ||| ||   | || |||||  | ||| | || ||||||||||||||
Db  85141 AGCGAGCACACGCCCACTA---AGGTGGCGTTAGTCGGCTACGATTATGTCGGCATGCGT 85085

Qy    121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180
          ||  | ||| |  |||| ||||||||    ||||  ||||| || ||||| |  ||||||
Db  85084 CCGACCATGCATGTCAAAGAAGGTGATAAGGTCACCAAAGGACAGGTGCTTTACGAAGAC 85025

Qy    181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240
          |||||     | |||||  | | ||| || || | |   ||||   |||  ||| |  |
Db  85024 AAAAAACGAGCAGGCGTCCTTTATACAGCACCAGTTGGCGGCACTGTCGTTGCTGTCAAT 84965

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttg------aaggcaacgac 294
```



Db 83884 CGTGATAACTTGACCCGTATTGAGCAGGAAGGCTAA 83849

RESULT 11
AAZ54034/c
ID AAZ54034 standard; DNA; 363 BP.
XX
AC AAZ54034;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 628 partial DNA sequence SEQ ID NO:2017.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR P-PSDB; AAY75272.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 7; Page 1003; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 363 BP; 74 A; 96 C; 110 G; 83 T; 0 other;

Query Match 24.2%; Score 324.6; DB 21; Length 363;
Best Local Similarity 93.4%; Pred. No. 4.2e-89;
Matches 339; Conservative 0; Mismatches 24; Indels 0; Gaps 0;



```
Qy    372 atcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccctgccgtagatgc 431
          ||| || || |||||||||||| |||||||||||||||||||||||| |||||||| |||||
Db    303 ATCCGGTTTGTGGACTGCGCTGCGCACCCGTCCGTTCAGCAAAATTCCTGCCGTCGATGC 244

Qy    432 cgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggctgccgaccctac 491
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 CGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCTGCCGACCCTAC 184

Qy    492 ggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggtattgagccgcct 551
          |||||| |||||||||||||||| || |||||||||||||||||||||||||||||  |
Db    183 GGTCATTATCAAAGAAGCCGCCGAGGATTTCAAACGCGGCCTGTTGGTATTGAGCCGTTT 124

Qy    552 gaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccgtctgaaaatgc 611
          ||||||||| |||||||||| ||||| ||||| ||||||||||||||||||||||||||
Db    123 GACCGAACGCAAAATCCATGTTTGTAAGGCAGCTGGCGCAGACGTGCCGTCTGAAAATGC 64

Qy    612 tgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttgagtggcacgca 671
          |||||| |||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db     63 TGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCTGCCGGTTTGAGTGGCACGCA 4

Qy    672 cat 674
          |||
Db      3 CAT 1
```

```
RESULT  12
AAZ54035/c
ID   AAZ54035 standard; DNA; 363 BP.
XX
AC   AAZ54035;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 628 partial DNA sequence SEQ ID NO:2019.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY75273.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 1003-1004; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 363 BP; 75 A; 105 C; 103 G; 80 T; 0 other;
```

```
  Query Match             22.4%;  Score 300.6;  DB 21;  Length 363;
  Best Local Similarity   89.3%;  Pred. No. 9.3e-82;
  Matches  324;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy    312 ctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgcaacctgattca 371
          |||||  || |||||| ||||||| || ||| ||||| |||||||||||||| ||||| ||
Db    363 CTACGCGCCCGAAGCGTTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGCAATCTGATCCA 304

Qy    372 atcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccctgccgtagatgc 431
          ||| || || |||||||||||| || |||||||||||||||||||||||||||||| |||||
Db    303 ATCCGGTTTGTGGACTGCGCTGCGTACCCGTCCGTTCAGCAAAATCCCTGCCGTCGATGC 244

Qy    432 cgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggctgccgaccctac 491
          ||||||||||||||||||||||||||||||||||||||||||||||||||| || ||||||
Db    243 CGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCGGCAGACCCTGT 184

Qy    492 ggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggtattgagccgcct 551
          |||  | |||||||||||||||||| || |||| ||| ||  || ||||||||||||  |
Db    183 GGTTGTGATCAAAGAAGCCGCCGAGGATTTCAGACGAGGTTTGCTGGTATTGAGCCGTTT 124

Qy    552 gaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccgtctgaaaatgc 611
          |||||| |||||||||||||||||| ||||| ||||||||||||||||||||||||||||
Db    123 GACCGAGCGTAAAATCCATGTGTGTAAGGCAGCTGGCGCAGACGTGCCGTCTGAAAATGC 64

Qy    612 tgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttgagtggcacgca 671
          |||||| |||||||||||||||| |||||||||||||| ||||| ||||||||||||||
Db     63 TGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCGGCCGGTTTGAGTGGCACGCA 4

Qy    672 cat 674
          |||
Db      3 CAT 1
```

```
RESULT  13
AAX91657
ID   AAX91657 standard; DNA; 1353 BP.
XX
AC   AAX91657;
XX
DT   25-AUG-1999  (first entry)
XX
DE   Porphorymonas gingivalis protein PG1 encoding DNA.
XX
KW   Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
KW   vaccine; antigenic; ds.
XX
OS   Porphorymonas gingivalis.
XX
PN   WO9929870-A1.
XX
PD   17-JUN-1999.
XX
PF   10-DEC-1998;   98WO-AU01023.
XX
PR   04-AUG-1998;   98AU-0005028.
PR   10-DEC-1997;   97AU-0000839.
PR   31-DEC-1997;   97AU-0001182.
```

PR 30-JAN-1998; 98AU-0001546.
PR 10-MAR-1998; 98AU-0002264.
PR 09-APR-1998; 98AU-0002911.
PR 23-APR-1998; 98AU-0003128.
PR 05-MAY-1998; 98AU-0003338.
PR 22-MAY-1998; 98AU-0003654.
PR 29-JUL-1998; 98AU-0004917.
XX
PA (CSLC-) CSL LTD.
XX
PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
PI Ross BC, Rothel LJ, Webb EA;
XX
DR WPI; 1999-385613/32.
DR P-PSDB; AAY34439.
XX
PT Antigenic Porphorymonas gingivalis peptides for preventing
PT gingivitis
XX
PS Claim 12; Page 179; 588pp; English.
XX
CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphorymonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphorymonas gingivalis. Probes can
CC be used to detect Porphorymonas gingivalis in standard hybridisation
CC assays. Porphorymonas gingivalis is involved in periodontal disease
CC especially gingivitis.
XX
SQ Sequence 1353 BP; 337 A; 333 C; 379 G; 304 T; 0 other;

Query Match 9.4%; Score 126.4; DB 20; Length 1353;
Best Local Similarity 46.7%; Pred. No. 3.9e-28;
Matches 631; Conservative 0; Mismatches 691; Indels 30; Gaps 6;

```
Qy      2 tgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcattt 61
          | || ||||  |||||||| ||    ||      |    || | ||||  ||  |  ||
Db     11 ttataaaaacaaaaaaaggccttgcacttaatctgaaaggaaaaccgctgcccgagatgc 70

Qy     62 atgacggcccggcc---attaccgaagtcgcgttgcttggcgaagaatatgtcggcatgc 118
            | ||  ||||||   | | |      |||| |  |   ||| || | ||  ||  |
Db     71 tggccgaaccggcccaaagtcctacttacgcggtcgtgcccgacgattttgaaggtgtta 130

Qy    119 gccccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaag 178
           ||||  | |||          || ||    ||       |||  ||  ||| |  |
Db    131 tccccaaggtgacggctcgtccgggggataaggtgcgtgccggctcagcactgatgcacc 190

Qy    179 acaaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattc 238
          ||||     ||||||   |    |||||    |||| |   ||| || |   |||  |
Db    191 acaaggcatatccggagatgaagtttacaagtccggttagcggcgaagtgatcgcggtga 250

Qy    239 accgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaa 298
          | || || |  ||||||    |   |    ||| | |       || ||     |||||
Db    251 atcgcggtgccaagcgcaaggtgttgagcatcgaggtgaaaccggacggactgaacgaat 310

Qy    299 tcgagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgcc 358
           |||||     |    ||    |      | | |      |      |||| |  | .
Db    311 acgagtcattccctgtcggggat------ccgtctgccctctctgccgaacagatcaagg 364

Qy    359 gcaacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatcc 418
              |  |       ||  | |||  |    ||      ||||| | |  || | |
Db    365 agcttttactgtcgagcggtatgtggggttttattaagcaacgtccttacgacatagtgg 424

Qy    419 ctgccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctgg 478
          || |    |||   |  ||   || ||| |   | | ||| |      ||    ||  |||
Db    425 ctacaccggatatagctccacgcgacatttatattactgccaactttactgcaccattgg 484

Qy    479 ctgccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttgg 538
          || | |||       |||||  |   || ||  |     | | |  |    | | |    |
Db    485 ctccggacttcgatttcatcgttcgaggagaagaacgcgccctgcagactgccatcgatg 544

Qy    539 tattgagccgcctgaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgc 598
            |||  |   || ||    | | ||  |  |||||  |         | ||  ||    ||
Db    545 ccttggccaaactcacgacaggaaaggtgtatgtgggcctgaagccgggttcatctctgg 604

Qy    599 cgtctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggct 658
            |   | |||||  |  |   | |||  |||          ||| || |||||  || ||
Db    605 gcttgcacaatgcagaaatcgtagaagtaca---------cggacctcatccggcaggta 655

Qy    659 tgagtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggacca 718
              |||  ||  ||  ||    |   ||||| ||     |    |||| |||||||||
Db    656 acgtgggcgtgctgatcaatcatacgaagccaatcaatcggggcgaaacggtgtggacgc 715

Qy    719 tcaattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaataccg 778
          ||||   |   ||| |||| |  ||||||||||  |  | || |||       ||
Db    716 tcaaggctaccgacctgatcgtgatcggacgtttcctgcttacgggcaaagccgatttta 775

Qy    779 agcgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgttt 838
            |  || ||||| ||  ||||    | |        |   |   | | ||||   |
Db    776 ccagaatgattgccatgaccggctcagacgctgcagctcacggatacgtccgtattatgc 835

Qy    839 tgggtgcgaaggtgtctcaacttaccgccggcgaattggtt------gacgcggacaacc 892
            ||||   || || | |        | |||||  | ||         ||  |  || | |
Db    836 cgggttgcaatgtctttgcttccttccccggccgactgacaataaaggaatctcacgagc 895

Qy    893 gcgtgatttccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgg 952
          | |||||    ||    ||  ||| ||||  ||        |    | |   ||   ||
Db    896 gtgtgatcgatggcaatgtgctgaccggtaagaagctctgcgagaaggagcctttcctgt 955

Qy    953 gacgctaccacaatcagatttccgttatcgaagaaggccgcagca---aagagctgttcg 1009
            |  |       | |||||  | || |||   ||||||  |       | || || ||||
Db    956 cagcccggtgtgaccagatcacggtgatccccgaaggcgacgatgtggacgaactcttcg 1015

Qy   1010 gctgggttgcgccgcagccggacaaatactccatcacgcgcaccactctcggccatttcc 1069
          | |||| ||| || |  |  ||  | ||| ||| |   |  |  | ||      || |
Db   1016 ggtgggctgcaccccgtctcgatcagtacagcatgagcagagcttatttctcttggttgc 1075

Qy   1070 ---taaaaaacaaactcttcaagttcacgacagccgtcaacggcggcgaccgcgccatgg 1126
               |||||||||   | |    ||     |    |||| || ||||| || || |||
Db   1076 aggggaaaaacaaagagtacgtactcgatgcccggatcaagggtggcgaacgtgctatga 1135

Qy   1127 taccgatcggcacttatgagcgcgtaatgccgttggacatcctgcctaccttgctttgc 1186
          |   || | |  ||||| |||||  | ||| ||||||||   ||    |  | |
Db   1136 tcatgagcaacgagtatgaccgcgttttcccgatggacatctatccggagtatttgctca 1195

Qy   1187 gcgatttaatcgtcggcgataccgacagcgcgcaggctttgggttgcttggaattggacg 1246
            | | | || |    ||| | |||||    | |||  || ||     |  ||| |||
Db   1196 aggctattatagcattcgacatcgacaagatggaggacttaggcatatatgaagtggctc 1255

Qy   1247 aagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgt 1306
            || ||| | ||    |||   || ||     |  |||   |||       |  |
Db   1256 cggaggactttgccacttgcgaatttgtggatacatccaagatcgagctgcagcgtatcg 1315

Qy   1307 tgcgcaaagtgctggaaaccattgagaaggaa 1338
          | ||| | |   |||| |   |  | ||||||
Db   1316 ttcgcgagggcttggatatgctctataaggaa 1347
```

RESULT 14
AAX91536
ID AAX91536 standard; DNA; 1362 BP.
XX
AC AAX91536;
XX
DT 25-AUG-1999 (first entry)
XX
DE Porphorymonas gingivalis protein PG1 ORF encoding DNA.
XX
KW Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
KW vaccine; antigenic; ds.
XX
OS Porphorymonas gingivalis.
XX

```
PN   WO9929870-A1.
XX
PD   17-JUN-1999.
XX
PF   10-DEC-1998;   98WO-AU01023.
XX
PR   04-AUG-1998;   98AU-0005028.
PR   10-DEC-1997;   97AU-0000839.
PR   31-DEC-1997;   97AU-0001182.
PR   30-JAN-1998;   98AU-0001546.
PR   10-MAR-1998;   98AU-0002264.
PR   09-APR-1998;   98AU-0002911.
PR   23-APR-1998;   98AU-0003128.
PR   05-MAY-1998;   98AU-0003338.
PR   22-MAY-1998;   98AU-0003654.
PR   29-JUL-1998;   98AU-0004917.
XX
PA   (CSLC-) CSL LTD.
XX
PI   Agius CT,  Barr IG,  Hocking DM,  Margetts MB,  Patterson MA;
PI   Ross BC,  Rothel LJ,  Webb EA;
XX
DR   WPI; 1999-385613/32.
DR   P-PSDB; AAY34318.
XX
PT   Antigenic Porphorymonas gingivalis peptides for preventing
PT   gingivitis
XX
PS   Claim 12; Page 96; 588pp; English.
XX
CC   AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC   Porphorymonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC   AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC   isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC   activity with a vaccine mechanism of action. The PG polypeptides can be
CC   used as vaccines especially against Porphorymonas gingivalis. Probes can
CC   be used to detect Porphorymonas gingivalis in standard hybridisation
CC   assays. Porphorymonas gingivalis is involved in periodontal disease
CC   especially gingivitis.
XX
SQ   Sequence 1362 BP; 337 A; 335 C; 381 G; 309 T; 0 other;
```

```
  Query Match             9.4%;  Score 126.4;  DB 20;  Length 1362;
  Best Local Similarity   46.7%;  Pred. No. 3.9e-28;
  Matches  631;  Conservative    0;  Mismatches  691;  Indels   30;  Gaps    6;

Qy       2 tgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcattt 61
           | || ||||  |||||||| ||    ||     |    || | ||||  ||  |  ||
Db      20 ttataaaaacaaaaaaaggccttgcacttaatctgaaaggaaaaccgctgcccgagatgc 79

Qy      62 atgacggcccggcc---attaccgaagtcgcgttgcttggcgaagaatatgtcggcatgc 118
             | ||  ||||||   | | |      |||| |  |   ||| || | ||  ||  |
Db      80 tggccgaaccggcccaaagtcctacttacgcggtcgtgcccgacgattttgaaggtgtta 139

Qy     119 gcccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaag 178
           ||||  | |||          || ||    ||       |||  ||  ||| |  |
Db     140 tccccaaggtgacggctcgtccgggggataaggtgcgtgccggctcagcactgatgcacc 199

Qy     179 acaaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattc 238
           ||||    ||||||   |   |||||   |||| |   ||| || |   |||  |
Db     200 acaaggcatatccggagatgaagtttacaagtccggttagcggcgaagtgatcgcggtga 259

Qy     239 accgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaa 298
           | || || |  ||||||    |   |    ||| | |       || ||     |||||
Db     260 atcgcggtgccaagcgcaaggtgttgagcatcgaggtgaaaccggacggactgaacgaat 319

Qy     299 tcgagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgcc 358
           |||||     |    ||     |      | | |     |      |||| | |
Db     320 acgagtcattccctgtcggggat------ccgtctgccctctctgccgaacagatcaagg 373

Qy     359 gcaacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatcc 418
             |  |       ||  | |||  |    ||      |||||  | |  || | |
```

```
Db     374 agcttttactgtcgagcggtatgtgggggttttattaagcaacgtccttacgacatagtgg 433

Qy     419 ctgccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctgg 478
           || |    |||    |  ||   || ||| |   | | ||| |     ||    ||  |||
Db     434 ctacaccggatatagctccacgcgacatttatattactgccaactttactgcaccattgg 493

Qy     479 ctgccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttgg 538
           || | |||       |||||  |   || ||  |     | | |   |   | | |    |
Db     494 ctccggacttcgatttcatcgttcgaggagaagaacgcgccctgcagactgccatcgatg 553

Qy     539 tattgagccgcctgaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgc 598
             |||  |   ||  ||   | | ||  |  ||||| |        | ||  ||   ||
Db     554 ccttggccaaactcacgacaggaaaggtgtatgtgggcctgaagccgggttcatctctgg 613

Qy     599 cgtctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggct 658
             |   | |||||  |  |   | |||  |||           ||| || ||||| || ||
Db     614 gcttgcacaatgcagaaatcgtagaagtaca---------cggacctcatccggcaggta 664

Qy     659 tgagtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggacca 718
             |||  ||  ||  ||   |  ||||| ||     |    |||| |||||||| 
Db     665 acgtgggcgtgctgatcaatcatacgaagccaatcaatcggggcgaaacggtgtggacgc 724

Qy     719 tcaattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaataccg 778
           ||||   |   ||| |||| |  ||||||||||||  |  | || |||       ||
Db     725 tcaaggctaccgacctgatcgtgatcggacgtttcctgcttacgggcaaagccgatttta 784

Qy     779 agcgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgttt 838
             |  || ||||| ||  ||||    | |         |   |   | | ||||   |
Db     785 ccagaatgattgccatgaccggctcagacgctgcagctcacggatacgtccgtattatgc 844

Qy     839 tgggtgcgaaggtgtctcaacttaccgccggcgaattggtt------gacgcggacaacc 892
           |||||   || || | |         | |||||  | ||           ||  |  || | |
Db     845 cgggttgcaatgtctttgcttccttccccggccgactgacaatataaggaatctcacgagc 904

Qy     893 gcgtgatttccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgg 952
           | |||||    ||    ||  ||| ||||  ||          |    | |   ||
Db     905 gtgtgatcgatggcaatgtgctgaccggtaagaagctctgcgagaaggagcctttcctgt 964

Qy     953 gacgctaccacaatcagatttccgttatcgaagaaggccgcagca---aagagctgttcg 1009
             | |       | |||||  | || |||   ||||||  |       | || || ||||
Db     965 cagcccggtgtgaccagatcacggtgatccccgaaggcgacgatgtggacgaactcttcg 1024

Qy    1010 gctgggttgcgccgcagccggacaaatactccatcacgcgcaccactctcggccatttcc 1069
           | |||| ||| || |  |  ||  | |||  ||| |   |  |  | ||      || |
Db    1025 ggtgggctgcaccccgtctcgatcagtacagcatgagcagagcttatttctcttggttgc 1084

Qy    1070 ---taaaaaacaaactcttcaagttcacgacagccgtcaacggcggcgaccgcgccatgg 1126
              |||||||||   | |    ||     |    |||| || ||||| || || |||
Db    1085 aggggaaaaacaaagagtacgtactcgatgcccggatcaagggtggcgaacgtgctatga 1144

Qy    1127 taccgatcggcacttatgagcgcgtaatgccgttggacatcctgcctaccttgcttttgc 1186
           |   || |  |   ||||| |||||  | ||| |||||||   ||    |  |  |
Db    1145 tcatgagcaacgagtatgaccgcgttttcccgatggacatctatccggagtatttgctca 1204

Qy    1187 gcgatttaatcgtcggcgataccgacagcgcgcaggctttgggttgcttggaattggacg 1246
           | | | || |    ||| | |||||    | |||  || ||     |  ||| |||
Db    1205 aggctattatagcattcgacatcgacaagatggaggacttaggcatatatgaagtggctc 1264

Qy    1247 aagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgt 1306
             || ||| | ||    |||  || ||     |   |||   |||        |  |
Db    1265 cggaggactttgccacttgcgaatttgtggatacatccaagatcgagctgcagcgtatcg 1324

Qy    1307 tgcgcaaagtgctggaaaccattgagaaggaa 1338
           | ||| | |  |||| |   |  | |||||||
Db    1325 ttcgcgagggcttggatatgctctataaggaa 1356
```

```
RESULT  15
AAD17186
ID   AAD17186 standard; DNA; 125401 BP.
XX
AC   AAD17186;
XX
```

```
DT   29-NOV-2001  (first entry)
XX
DE   Streptomyces noursei nystatin PKS gene cluster DNA.
XX
KW   Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;
KW   antifungal; antibiotic; ds.
XX
OS   Streptomyces noursei.
XX
FH   Key             Location/Qualifiers
FT   CDS             6337..34771
FT                   /*tag= a
FT                   /product= "NysI complete protein"
FT   CDS             34792..51099
FT                   /*tag= b
FT                   /product= "NysJ protein"
FT   CDS             51155..57355
FT                   /*tag= c
FT                   /product= "NysK protein"
FT   CDS             57503..58687
FT                   /*tag= d
FT                   /product= "NysL protein"
FT   CDS             complement (58786..58980)
FT                   /*tag= e
FT                   /product= "NysM protein"
FT                   /note= "CDS does not include start codon"
FT   CDS             complement (59045..60241)
FT                   /*tag= f
FT                   /product= "NysN protein"
FT                   /note= "CDS does not include start codon"
FT   CDS             complement (60238..61296)
FT                   /*tag= g
FT                   /product= "NysD2 complete protein"
FT   CDS             120628..121308
FT                   /*tag= h
FT                   /product= "NysR4 (long) protein"
XX
PN   WO200159126-A2.
XX
PD   16-AUG-2001.
XX
PF   08-FEB-2001; 2001WO-GB00509.
XX
PR   08-FEB-2000; 2000GB-0002840.
PR   10-APR-2000; 2000GB-0008786.
PR   14-APR-2000; 2000GB-0009387.
XX
PA   (UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
PA   (SNTF ) SINTEF STIFTELSEN IND TEK FORSK.
PA   (ALPH-) ALPHARMA AS.
PA   (SINV-) SINVENT AS.
PA   (DZIE/) DZIEGLEWSKA H.
PA   (ZOTC/) ZOTCHEV S B.
PA   (SEKU/) SEKUROVA O N.
PA   (FJAE/) FJAERVIK E.
PA   (BRAU/) BRAUTASET T.
PA   (STRO/) STROM A R.
XX
PI   Zotchev SB,  Sekurova ON,  Fjaervik E,  Brautaset T,  Strom AR;
PI   Valla S,  Ellingsen TE,  Sletta H,  Gulliksen O;
XX
DR   WPI; 2001-557614/62.
DR   P-PSDB; AAE10143, AAE10144, AAE10145, AAE10146, AAE10147, AAE10148,
DR   AAE10149, AAE10150.
XX
PT   New nystatin polyketide synthase polynucleotides and polypeptides,
PT   useful as antibiotics and antifungals -
XX
PS   Claim 1; Page 188-254; 266pp; English.
XX
CC   The present invention relates to the cloning and sequencing of the gene
CC   cluster encoding a modular type I polyketide synthase (PKS) enzyme
CC   involved in the biosynthesis of the macrolide antibiotic nystatin.
CC   The nystatin PKS is useful as antifungal antibiotics. The present
CC   sequence is a Streptomyces noursei nystatin PKS gene cluster DNA.
XX
SQ   Sequence 125401 BP; 15664 A; 49692 C; 42871 G; 17174 T; 0 other;


  Query Match             2.9%;  Score 38.8;  DB 22;  Length 125401;
  Best Local Similarity   48.6%;  Pred. No. 2.3;
  Matches  106;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy    1104 caacggcggcgaccgcgccatggtaccgatcggcacttatgagcgcgtaatgccgttgga 1163
           || || |||||| || || |   |   |  |||||||      || ||    || |
Db   50040 caccgccggcgtcctcgacgacgccaccgtcggcaccctgacgcccgagcggctggacac 50099

Qy    1164 catcctgcctaccttgcttttgcgcgatttaatcgtcggcgataccgacagcgcgcaggc 1223
           | ||||||   ||  |       ||  |    | |   |||  ||  | |||  |  |
Db   50100 cgtcctgcgcgccaaggcggacgccgcctggcacctgcacgacgccacccgcgaccgcga 50159

Qy    1224 tttgggttgcttggaattggacgaagaagacctcgctttgtgcagcttcgtctgcccggg 1283
             |||   | ||  |   || ||      | |  || | |   | ||  |  | ||| |||
Db   50160 cctggcagggttcgtgctgtactcctcggtcgccggtgtcaccggcggccccggccaggg 50219

Qy    1284 caaatacgaatacggcccgctgttgcgcaaagtgctgg 1321
           ||| ||||    ||||     ||| | | | | ||| |
Db   50220 caactacgccgccggcaacacgttcctcgacgcgctcg 50257


RESULT  16
AAZ32020
ID   AAZ32020 standard; DNA; 38734 BP.
XX
AC   AAZ32020;
XX
DT   10-JAN-2000  (first entry)
XX
DE   Human METH1 related EST AL021529.
XX
KW   Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW   cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW   angiogenesis inhibitor; abnormal wound healing; inflammation;
KW   rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW   diabetic retinopathy; macula degeneration; haemangioma; detection;
KW   arterial-venous malformation; immune deficiency; ss.
XX
OS   Homo sapiens.
XX
PN   WO9937660-A1.
XX
PD   29-JUL-1999.
XX
PF   22-JAN-1999;   99WO-US01313.
XX
PR   23-JAN-1998;   98US-0072298.
PR   28-AUG-1998;   98US-0098539.
XX
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM;
XX
DR   WPI; 1999-590684/50.
XX
PT   New isolated metalloprotease thrombospondin polypeptides, useful for
PT   treating hyperproliferative disorders, cancers or autoimmune disorders
PT   -
XX
PS   Disclosure; Page 296-321; 457pp; English.
XX
CC   AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC   metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC   respectively. METH1 and METH2 have been found to be potent inhibitors of
CC   angiogenesis both in vitro and in vivo. They can be used for treating
```

```
CC   cancer and other disorders related to angiogenesis including abnormal
CC   wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC   endometrial bleeding disorders, diabetic retinopathy, some forms of
CC   macula degeneration, haemangiomas, and arterial-venous malformations.
CC   They may be useful in treating deficiencies or disorders of the immune
CC   system, by activating or inhibiting the proliferation, differentiation,
CC   or mobilisation (chemotaxis) of immune cells. The etiology of these
CC   immune deficiencies or disorders may be genetic, somatic, such as
CC   cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC   toxins), or infectious. They can also be used to treat inflammatory
CC   conditions, both chronic and acute conditions. The products can also be
CC   used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to
CC   AAY49511 represent sequences given in the exemplification of the present
CC   invention.
XX
SQ   Sequence 38734 BP; 6142 A; 13140 C; 13585 G; 5867 T; 0 other;

  Query Match             2.8%;  Score 38.2;  DB 20;  Length 38734;
  Best Local Similarity   48.0%;  Pred. No. 2;
  Matches  109;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy    1116 ccgcgccatggtaccgatcggcacttatgagcgcgtaatgccgttggacatcctgcctac 1175
           | ||| |||  |  || | | | |         |||   |  || ||| | || | | ||
Db    4519 ctgcggcatcatcgcgctggccgccgcgaccgacgtgcggatgtcggagaacccggccac 4578

Qy    1176 cttgcttttgcgcgatttaatcgtcggcgataccgacagcgcgcaggctttgggttgctt 1235
           |  || || | | |     |||  | ||   | |||  ||  |||     ||    ||
Db    4579 cgacctcttccacaacggcgtcgcggtcggcgcggactacgtccagcacccggtgatctc 4638

Qy    1236 ggaattggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaata 1295
           | |  | | ||| |  | | |||    | |||||||| | | |   |       || |
Db    4639 gcaggtcgccgaggcggtcttcggcgagggcagcttcctgttcatcgtgctggccgcagc 4698

Qy    1296 cggcccgctgttgcgcaaagtgctggaaaccattgagaaggaaggct 1342
           |  | |||| | | |    |    |  | ||  | | |  | ||||
Db    4699 caccgcgctggtcctcttcctcgccgccaacaccgcgtacaacggct 4745


RESULT  17
AAC90077
ID   AAC90077 standard; DNA; 38734 BP.
XX
AC   AAC90077;
XX
DT   19-MAR-2001  (first entry)
XX
DE   AL021529 cDNA clone.
XX
KW   METH; metalloprotease; thrombospondin; angiogenesis inhibition;
KW   cancer therapy; benign tumour; ocular angiogenic disease;
KW   rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW   vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW   scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW   coronary collateral; cerebral collateral; arteriovenous malformation;
KW   ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;
KW   plaque neovascularisation; telangiectasia; haemophiliac joint; EST;
KW   angiofibroma; fibromuscular dysplasia; expressed sequence tag;
KW   Crohn's disease; atherosclerosis; birth control; ss.
XX
OS   Unidentified.
XX
PN   WO200071577-A1.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-US14462.
XX
PR   25-MAY-1999;   99US-0318208.
PR   20-JUL-1999;   99US-0144882.
PR   10-AUG-1999;   99US-0147823.
PR   13-AUG-1999;   99US-0373658.
PR   22-DEC-1999;   99US-0171503.
PR   22-FEB-2000; 2000US-0183792.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
PA   (JONA/) JONAK Z L.
PA   (TRUL/) TRULLI S H.
PA   (FORN/) FORNWALD J A.
PA   (TERR/) TERRETT J A.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM,  Jonak ZL,  Trulli SH;
PI   Fornwald JA,  Terrett JA;
XX
DR   WPI; 2001-025136/03.
XX
PT   METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT   inhibit angiogensis in the treatment of disorders such as cancer,
PT   rheumatoid arthritis and psoriasis -
XX
PS   Claim 7; Pages 597-622; 768pp; English.
XX
CC   The present invention relates to human METH1 and METH2, (ME for
CC   metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
CC   The present sequence is an expressed sequence tag (EST) for METH. METH
CC   can be used for inhibiting angiogenesis in an individual, and for
CC   treating cancer, benign tumours, an ocular angiogenic disease,
CC   rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC   vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
CC   scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
CC   coronary collaterals, cerebral collaterals, arteriovenous malformations,
CC   ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
CC   neovascularisation, telangiectasia, haemophiliac joints, angiofibroma,
CC   fibromuscular dysplasia, wound granulation, Crohn's disease or
CC   atherosclerosis. METH can also be used in birth control. METH can also
CC   be used in diagnostic methods for the prognosis of cancer.
XX
SQ   Sequence 38734 BP; 6142 A; 13138 C; 13586 G; 5868 T; 0 other;

  Query Match             2.8%;  Score 38.2;  DB 22;  Length 38734;
  Best Local Similarity   48.0%;  Pred. No. 2;
  Matches  109;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy    1116 ccgcgccatggtaccgatcggcacttatgagcgcgtaatgccgttggacatcctgcctac 1175
           | ||| |||  |  || | | | |         |||   |  || ||| | || | | ||
Db    4519 ctgcggcatcatcgcgctggccgccgcgaccgacgtgcggatgtcggagaacccggccac 4578

Qy    1176 cttgcttttgcgcgatttaatcgtcggcgataccgacagcgcgcaggctttgggttgctt 1235
           |  || || | | |     |||  | ||   | |||  ||  |||     ||    ||
Db    4579 cgacctcttccacaacggcgtcgcggtcggcgcggactacgtccagcacccggtgatctc 4638

Qy    1236 ggaattggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaata 1295
           | |  | | ||| |  | | |||    | |||||||| | | |   |       || |
Db    4639 gcaggtcgccgaggcggtcttcggcgagggcagcttcctgttcatcgtgctggccgcagc 4698

Qy    1296 cggcccgctgttgcgcaaagtgctggaaaccattgagaaggaaggct 1342
           |  | |||| | | |    |    |  | ||  | | |  | ||||
Db    4699 caccgcgctggtcctcttcctcgccgccaacaccgcgtacaacggct 4745


RESULT  18
AAD13450
ID   AAD13450 standard; DNA; 5036 BP.
XX
AC   AAD13450;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Mycobacterium tuberculosis cosmid MTY13E10 DNA.
XX
```

```
KW   Isoniazid induced gene; ini; antituberculosis agent; isoniazid; INH;
KW   ethambutol; EMB; infection; tuberculosis; polio; leprosy; malaria;
KW   tetanus; diphtheria; influenza; measles; mumps; hepatitis; rabies;
KW   drug screening; tuberculostatic; human immunodeficiency virus; HIV;
KW   protozoacide; hepatotropic; virucide; antiinflammatory; vaccine;
KW   cosmid MTY13E10; ds.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             160..1599
FT                   /*tag= a
FT                   /product= "M. tuberculosis iniB protein, Rv0341"
FT   CDS             1636..3558
FT                   /*tag= b
FT                   /product= "M. tuberculosis iniA protein, Rv0342"
FT   CDS             3555..5036
FT                   /*tag= c
FT                   /product= "M. tuberculosis iniC protein, Rv0343"
FT                   /transl_except= (pos: 3555..3557. aa: Met)
XX
PN   US6268201-B1.
XX
PD   31-JUL-2001.
XX
PF   23-OCT-1998;   98US-0177349.
XX
PR   23-OCT-1998;   98US-0177349.
XX
PA   (YESH ) UNIV YESHIVA EINSTEIN COLLEGE.
XX
PI   Alland D,  Bloom BR,  Jacobs WR;
XX
DR   WPI; 2001-496015/54.
DR   P-PSDB; AAE07166, AAE07167, AAE07168.
XX
PT   Novel isoniazid induced gene B obtained from Mycobacterium
PT   tuberculosis, useful as vaccine for preventing and treating
PT   tuberculosis, human immunodeficiency viral infection, leprosy, malaria,
PT   tetanus and diphtheria -
XX
PS   Claim 1; Fig 5; 26pp; English.
XX
CC   The invention relates to the identification, cloning, sequencing and
CC   characterisation of the isoniazid induced gene A (iniA), iniB and iniC
CC   of mycobacteria which are induced by a broad class of antibiotics that
CC   act by inhibiting cell wall biosynthesis, including the first line
CC   antituberculosis agents, isoniazid (INH) and ethambutol (EMB). iniB
CC   gene is useful as vaccine or to improve existing vaccines, for treating
CC   mycobacterial infections, caused by mycobacteria such as M. bovis,
CC   M. tuberculosis, and M. leprae. These vaccines can be used to prevent
CC   and treat a wide variety of diseases, including tuberculosis, human
CC   immunodeficiency viral infection, polio, leprosy, malaria, tetanus,
CC   diphtheria, influenza, measles, mumps, hepatitis and rabies. Determining
CC   if a mutated iniB gene exists in a mycobacterial strain is useful for
CC   assessing the susceptibility of the mycobacterium to EMB. Vector
CC   construct comprising the nucleotide sequence of an iniB promoter from
CC   M. tuberculosis is useful for screening drugs or compounds to determine
CC   if the drug or compound is effective against M. tuberculosis. The
CC   present DNA sequence is Mycobacterium tuberculosis cosmid MTY13E10
CC   encoding iniA, iniB and iniC proteins.
XX
SQ   Sequence 5036 BP; 805 A; 1590 C; 1762 G; 879 T; 0 other;

  Query Match             2.7%;  Score 35.8;  DB 22;  Length 5036;
  Best Local Similarity   54.1%;  Pred. No. 4.1;
  Matches   73;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy     415 atccctgccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
           |||| |||   |     ||||| ||||| ||| |||| ||   ||||   |       |
Db    4458 atccgtgcgcaattgctcgagcggttcggcatgttcggcatccggatgtcgattgccgtg 4517

Qy     475 ctggctgccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctg 534
           ||||| |||| |   || |     | |      |||||||||    |  | ||| ||  |
Db    4518 ctggcggccggcgtgaccgattcgaccgggctggccgccgaactgctggagcgcagcggg 4577

Qy     535 ttggtattgagccgc 549
           ||||   |   |||
Db    4578 ctggtggcgctgcgc 4592


RESULT  19
AAI99683
ID   AAI99683 standard; DNA; 4403765 BP.
XX
AC   AAI99683;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ -
XX
PS   Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;

  Query Match             2.7%;  Score 35.8;  DB 22;  Length 4403765;
  Best Local Similarity   54.1%;  Pred. No. 95;
  Matches   73;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy     415 atccctgccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
           |||| |||   |     ||||| ||||| ||| |||| ||   ||||   |       |
Db 413713 atccgtgcgcaattgctcgagcggttcggcatgttcggcatccggatgtcgattgccgtg 413772

Qy     475 ctggctgccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctg 534
           ||||| |||| |   || |     | |      |||||||||    |  | ||| ||  |
Db 413773 ctggcggccggcgtgaccgattcgaccgggctggccgccgaactgctggagcgcagcggg 413832
```

```
Qy     535 ttggtattgagccgc 549
           ||||   |   |||
Db 413833 ctggtggcgctgcgc 413847


RESULT  20
AAI99682
ID   AAI99682 standard; DNA; 4411529 BP.
XX
AC   AAI99682;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ  -
XX
PS   Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;


  Query Match             2.7%;  Score 35.8;  DB 22;  Length 4411529;
  Best Local Similarity   54.1%;  Pred. No. 95;
  Matches   73;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy     415 atccctgccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
           |||| |||   |     ||||| ||||| ||| |||| ||    ||||   |       |
Db 413658 atccgtgcgcaattgctcgagcggttcggcatgttcggcatccggatgtcgattgccgtg 413717

Qy     475 ctggctgccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctg 534
           ||||| |||| |   || |     | |      |||||||||    |  | ||| ||  |
Db 413718 ctggcggccggcgtgaccgattcgaccgggctggccgccgaactgctggagcgcagcggg 413777

Qy     535 ttggtattgagccgc 549
           ||||   |   |||
Db 413778 ctggtggcgctgcgc 413792


RESULT  21
AAX84332/c
ID   AAX84332 standard; DNA; 5059 BP.
XX
AC   AAX84332;
XX
DT   08-SEP-1999  (first entry)
XX
DE   Stealth virus nucleic acid clone, SEQ ID NO: 24.
XX
KW   Stealth virus; detection; diagnosis; infection; ss.
XX
OS   Stealth virus.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 3605
FT                   /*tag=  a
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3610
FT                   /*tag=  b
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3615
FT                   /*tag=  c
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3630
FT                   /*tag=  d
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3631
FT                   /*tag=  e
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3638
FT                   /*tag=  f
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3641
FT                   /*tag=  g
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3642
FT                   /*tag=  h
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3657
FT                   /*tag=  i
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3659
FT                   /*tag=  j
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3681
FT                   /*tag=  k
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
```

```
FT   misc_difference 3698
FT                   /*tag= l
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3703
FT                   /*tag= m
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3708
FT                   /*tag= n
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3713
FT                   /*tag= o
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3720
FT                   /*tag= p
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3732
FT                   /*tag= q
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3733
FT                   /*tag= r
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3740
FT                   /*tag= s
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3747
FT                   /*tag= t
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3750
FT                   /*tag= u
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3751
FT                   /*tag= v
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3766
FT                   /*tag= w
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3778
FT                   /*tag= x
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3792
FT                   /*tag= y
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3799
FT                   /*tag= z
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3800
FT                   /*tag= aa
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3812
FT                   /*tag= ab
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3813
FT                   /*tag= ac
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3818
FT                   /*tag= ad
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3819
FT                   /*tag= ae
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3824
FT                   /*tag= af
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3827
FT                   /*tag= ag
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3841
FT                   /*tag= ah
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3853
FT                   /*tag= ai
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3857
FT                   /*tag= aj
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3860
FT                   /*tag= ak
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3863
FT                   /*tag= al

  Query Match              2.6%;  Score 35.6;  DB 20;  Length 5059;
  Best Local Similarity   15.9%;  Pred. No. 4.7;
  Matches   43;  Conservative   98;  Mismatches  129;  Indels    0;  Gaps    0;

Qy   1004 tgttcggctgggttgcgccgcagccggacaaatactccatcacgcgcaccactctcggcc 1063
          |::: : :|:: ::::|: | |:: : ::   :  |  :| |  :: |  |    :||
Db    598 TKYHGSYBTRVYDWSSGYAGNASHTSTHSYSTWNATATHTDATASSNAGNANATNYGGNT 539

Qy   1064 atttcctaaaaaacaaactcttcaagttcacgacagccgtcaacggcggcgaccgcgcca 1123
             |:  |      :::: |:  : : |::  : :|: :: :   :  |  : |   : :|
Db    538 GSTBAGTNCNNTKYRVHNTSNMYBHATYHNTSNMARGBSSYSNKKNAGHNSNCANTRAVA 479

Qy   1124 tggtaccgatcggcacttatgagcgcgtaatgccgttggacatcctgcctaccttgcttt 1183
```

```
          |:|    : :     |: ::: : :  :  :|:  :: :  :| | :   :::  :  |:
Db     478 TKGASDNSSYDANATCYNWHSTSKSNTDGKDGSDCYDASSNHTGCNKKDAWYYANSTRTK 419

Qy    1184 tgcgcgatttaatcgtcggcgataccgacagcgcgcaggctttgggttgcttggaattgg 1243
           ::: |:   |:  :  : |: | || : : :::   |: ::  ::: ,, |:| :  |
Db     418 SSYBNTAYGGNAKGNKKNSAGHGANCGGSYBDSYBRGNGHNKKNTKDDAWNAGNANHGNG 359

Qy    1244 acgaagaagacctcgctttgtgcagcttcg 1273
           :  :|:: :|: | :  :|:|:   :|:::
Db     358 RTYDAKDGRAYATKKASYTSTKTYHYTKSK 329
```

RESULT 22
AAI99683/c
ID AAI99683 standard; DNA; 4403765 BP.
XX
AC AAI99683;
XX
DT 15-JAN-2002 (first entry)
XX
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US6294328-B1.
XX
PD 25-SEP-2001.
XX
PF 24-JUN-1998; 98US-0103840.
XX
PR 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX
DR WPI; 2001-647261/74.
XX
PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ -
XX
PS Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
CC The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;

Query Match 2.6%; Score 35.6; DB 22; Length 4403765;
Best Local Similarity 57.0%; Pred. No. 1.1e+02;
Matches 65; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```
Qy     777 cgagcgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgt 836
           || |  ||||| || || |||| | |||    | ||   |||   ||  ||||  |||
Db 4270822 CGCGGAGGTGGTGGCGTTCGGCGACATGCCCAACGACGTGCCGATGCTGCTGCGGGCCGG 4270763

Qy     837 tttgggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaa 890
           |||||||| |  | ||  | |        ||| ||   |||  | ||| ||| |
Db 4270762 TTTGGGTGTGGCGATGGGTAACGCGCATCCCGACGCGCTGGCCGTCGCCGACGA 4270709
```

RESULT 23
AAI99682/c
ID AAI99682 standard; DNA; 4411529 BP.
XX
AC AAI99682;
XX
DT 15-JAN-2002 (first entry)
XX
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US6294328-B1.
XX
PD 25-SEP-2001.
XX
PF 24-JUN-1998; 98US-0103840.
XX
PR 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX
DR WPI; 2001-647261/74.
XX
PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ -
XX
PS Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
CC The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;

Query Match 2.6%; Score 35.6; DB 22; Length 4411529;
Best Local Similarity 57.0%; Pred. No. 1.1e+02;
Matches 65; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```
Qy     777 cgagcgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgt 836
           || |  ||||| || || |||| | |||    | ||   |||   ||  ||||  |||
Db 4278565 CGCGGAGGTGGTGGCGTTCGGCGACATGCCCAACGACGTGCCGATGCTGCTGCGGGCCGG 4278506

Qy     837 tttgggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaa 890
           |||||||| |  | ||  | |        ||| ||   |||  | ||| ||| |
Db 4278505 TTTGGGTGTGGCGATGGGTAACGCGCATCCCGACGCGCTGGCCGTCGCCGACGA 4278452
```

RESULT 24
AAS57519

```
ID   AAS57519 standard; cDNA; 296 BP.
XX
AC   AAS57519;
XX
DT   13-FEB-2002  (first entry)
XX
DE   cDNA #195 encoding portion of a human colon tumour protein.
XX
KW   Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200173027-A2.
XX
PD   04-OCT-2001.
XX
PF   22-MAR-2001; 2001WO-US09246.
XX
PR   24-MAR-2000; 2000US-191597P.
PR   04-MAY-2000; 2000US-202024P.
PR   05-MAY-2000; 2000US-202189P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Meagher MJ,  Xu J,  King GE;
XX
DR   WPI; 2001-611627/70.
XX
PT   New colon tumour proteins and related nucleic acid, useful for
PT   treatment, prevention, diagnosis and monitoring of cancer -
XX
PS   Claim 4; Page 90; 299pp; English.
XX
CC   Th present invention relates to the isolation of novel cDNA sequences
CC   encoding for at least an immunogenic portion of human colon tumour
CC   proteins. The sequences of the invention are useful in pharmaceutical
CC   compositions and vaccines for the prevention and treatment of cancers
CC   such as colon cancer. They are also useful for the diagnosis and
CC   monitoring of such cancers. Antibodies to the colon tumour proteins
CC   and antigen presenting cells that express polynucleotides encoding
CC   colon tumour proteins can be used to inhibit the development of
CC   cancers. T-cells that react specifically with colon tumour proteins
CC   are useful for removing tumour cells from samples (e.g. blood) and
CC   for cancer treatment. The polynucleotides sequences are also useful in
CC   gene therapy. AAS57325-AAS58880 represent the cDNA sequences of the
CC   invention that encode for portions of human colon tumour proteins.
XX
SQ   Sequence 296 BP; 66 A; 86 C; 88 G; 51 T; 5 other;
```

```
  Query Match             2.6%;  Score 35.4;  DB 23;  Length 296;
  Best Local Similarity   54.8%;  Pred. No. 1.4;
  Matches   69;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy   1215 cgcgcaggctttgggttgcttggaattggacgaagaagacctcgctttgtgcagcttcgt 1274
           || || | |    ||      |   |  || ||  || ||    || |   || || |
Db     50 cgagcggccgcccgggcaggtctcanggggccgccgaggatggggcctgcggctgcctgcc 109

Qy   1275 ctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctggaaaccattgagaa 1334
           |  |||||||| || |||  |  |||  | |  | |||| | ||||    ||||| |||
Db    110 caacccgggcacattcgaggagtgccaccggaagtgcaaggagctgtttcccattcagat 169

Qy   1335 ggaagg 1340
           ||| ||
Db    170 ggaggg 175
```

```
RESULT  25
AAA10594
ID   AAA10594 standard; DNA; 10732 BP.
XX
AC   AAA10594;
XX
DT   29-JUN-2000  (first entry)
XX
DE   Gene encoding a subunit of cellulose synthase.
XX
KW   Cellulose synthase; cellulose production; increase yield; ds.
XX
OS   Vigna angularis.
XX
PN   JP2000060568-A.
XX
PD   29-FEB-2000.
XX
PF   26-AUG-1998;   98JP-0239998.
XX
PR   26-AUG-1998;   98JP-0239998.
XX
PA   (MIZU/) MIZUNO K.
PA   (OJIP ) OJI PAPER CO.
XX
DR   WPI; 2000-342371/30.
DR   P-PSDB; AAY85179.
XX
PT   A gene encoding a cellulose synthetic equipment - for the improvement
PT   in the amount of cellulose synthesised in a plant body
XX
PS   Claim 2; Page 14-21; 32pp; Japanese.
XX
CC   This sequence represents a gene encoding a subunit of the cellulose
CC   synthase complex of Vigna angularis. The invention relates to subunits of
CC   cellulose synthetic equipment, that can be used to increase the amount of
CC   cellulose synthesised by a plant. The proteins and genes encoding them
CC   can also be used to improve the properties of the cellulose being
CC   produced by a plant.
XX
SQ   Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 2251 other;
```

```
  Query Match             2.6%;  Score 35.4;  DB 21;  Length 10732;
  Best Local Similarity   16.7%;  Pred. No. 7.8;
  Matches   37;  Conservative   95;  Mismatches   89;  Indels    0;  Gaps    0;

Qy     98 gcgaagaatatgtcggcatgcgcccctcgatgaaaatcaaggaaggtgaagccgtcaaaa 157
          :: : :: : ::::|:::  : :: :   | ::| |  :  :|  : |: |::  :|:|
Db   9582 shargsrgysysysgrsrcysyssrvagyaasrasasnmtasashvagmtgysaavavas 9641

Qy    158 aaggccaagtgctgtttgaagacaaaaagaatccgggcgtagtatttactgcgccggctt 217
          :: |::|| ::  :   :||::: ||:| ::  : ::::     :|::: |:: :::
Db   9642 rvagysaathrgnsrrsraasrysaavathrgasnsrsrgthrgythrysgsrthrrggg 9701

Qy    218 caggcaaaatcgccgctattcaccgtggcgaaaagcgcgtacttcagtcagtcgtgattg 277
          : :| :|:| ::: | |: |::       : :|:  ::  ::|  |::: : ::
Db   9702 yysgrvarashssrgnthrthrgthrthrrgvavaasmtgrmtsrashssrassryssry 9761

Qy    278 ccgttgaaggcaacgacgaaatcgagttcgaacgctacgta 318
          : : | :  :::|:  :|  : : :|:| | |::  |::|:
Db   9762 sasnthrcyssrashysgythrvargytrgnasvavaysth 9802
```

```
RESULT  26
AAS66442
ID   AAS66442 standard; cDNA; 1938 BP.
XX
AC   AAS66442;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #2246.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
```

```
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG02255.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 2246; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1938 BP; 500 A; 465 C; 516 G; 457 T; 0 other;


  Query Match             2.6%;  Score 35.2;  DB 23;  Length 1938;
  Best Local Similarity   51.2%;  Pred. No. 3.9;
  Matches   82;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy     151 gtcaaaaaaggccaagtgctgtttgaagacaaaaagaatccgggcgtagtatttactgcg 210
           || ||| ||| | || |          | | |   | | |  |||||| |  |   ||||
Db    1048 gttaaagaagtcgaaatcgaaacgcccggcgacctggaacacggcgtatttatcggtgcg 1107

Qy     211 ccggcttcaggcaaaatcgccgctattcaccgtggcgaaaagcgcgtacttcagtcagtc 270
           |||     |||||||||||| |       || |  ||  |    | |  |    ||  |
Db    1108 gcgggcgatggcaaaatcgcctcagcatcgcgggcagattatgcgggagctgctgcacgc 1167

Qy     271 gtgattgccgttgaaggcaacgacgaaatcgagttcgaac 310
           ||||||  ||  |  ||  |||| |  |  |  | |||||
Db    1168 gtgattagcgaagccggtcacgaaggcaaggtttacgaac 1207


RESULT  27
AAS86095
ID   AAS86095 standard; cDNA; 323 BP.
XX
AC   AAS86095;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #21899.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG21908.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 21899; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 323 BP; 95 A; 86 C; 84 G; 58 T; 0 other;


  Query Match             2.6%;  Score 35;  DB 23;  Length 323;
  Best Local Similarity   59.6%;  Pred. No. 1.9;
  Matches   59;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy    1242 ggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggccc 1301
           || ||  || ||    || |   || || |   |  |||||||| || |||  |  |||
Db      37 ggccgccgaggatggggcctgcggctgcctgcccaacccgggcacattcgaggagtgcca 96

Qy    1302 gctgttgcgcaaagtgctggaaaccattgagaaggaagg 1340
             |  |  |||| | ||||    ||||| ||| ||| ||
Db      97 ccggaagtgcaaggagctgtttcccattcagatggaggg 135


RESULT  28
AAF25080/c
ID   AAF25080 standard; DNA; 697 BP.
XX
AC   AAF25080;
```

```
XX
DT   30-APR-2001  (first entry)
XX
DE   Nucleotide sequence of a M. vaccae immunogenic epitope from DNA9A.
XX
KW   Epitope; antigen; cytokine production; immune disorder; tuberculosis;
KW   cancer; mycobacterial infection; TH1 immune response; vaccine; ss.
XX
OS   Mycobacterium vaccae.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..225
FT                   /*tag= a
FT                   /note= "no termination sequence given"
XX
PN   WO200104140-A1.
XX
PD   18-JAN-2001.
XX
PF   10-JUL-2000; 2000WO-NZ00121.
XX
PR   12-JUL-1999;   99US-0351348.
PR   29-NOV-1999;   99US-0450072.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Delcayre A;
XX
DR   WPI; 2001-168411/17.
DR   P-PSDB; AAB31649.
XX
PT   Novel polypeptides comprising  immunogenic epitopes of Mycobacterium
PT   vaccae, useful for treating mycobacterial infections, immune disorders
PT   and cancers  -
XX
PS   Claim 4; Page 59; 80pp; English.
XX
CC   The specification describes an immunogenic epitope of a Mycobacterium
CC   vaccae antigen. The epitope is a stimulator of cytokine production.
CC   The epitopes are useful for the treatment of immune disorders,
CC   infectious diseases, especially tuberculosis, and cancer. They are
CC   also useful for treatment of other mycobacterial infections such as
CC   those caused by Mycobacterium avium. The epitopes are especially useful
CC   for inducing TH1 immune responses, and for producing vaccines. The
CC   present sequence encodes a M. vaccae epitope of the invention.
XX
SQ   Sequence 697 BP; 89 A; 253 C; 253 G; 102 T; 0 other;

  Query Match             2.6%;  Score 35;  DB 22;  Length 697;
  Best Local Similarity   46.5%;  Pred. No. 2.7;
  Matches  113;  Conservative    0;  Mismatches  130;  Indels    0;  Gaps    0;

Qy    277 gccgttgaaggcaacgacgaaatcgagttcgaacgctacgtacctgaagcgctggcaaaa 336
          |||||    |||  |||||    || || ||    |   || |  || |||||||||
Db    447 GCCGTGCGCGGCGCCGACGCCGACGCGTACGCCGCCGCGGTTCGGGATGCGCTGGCACTG 388

Qy    337 ttgagcagcgaaaaagtgcgccgcaacctgattcaatcaggcttatggactgcgcttcgc 396
          |      |   | |||      ||   | |     | | |        ||   | |
Db    387 ATCGAAGAGGCCGGACAGCGGGCGGTCCGCCTGCCCGGCGACCTGGCCGAGCCGGCGCAC 328

Qy    397 acccgtccgttcagcaaaatccctgccgtagatgccgagccgttcgccatcttcgtcaat 456
            |||  || |   | |   | | |    | |     |       | | |||| |||||||
Db    327 TGCCGCGCGGTGGTCGAGGACACGGTGAAACAGCTTGGCGGACTGGACATCGTCGTCAAC 268

Qy    457 gcgatggacaccaatccgctggctgccgaccctacggtcatcatcaaagaagccgccgaa 516
             |||    | |  ||| |   | |||||  || |  |||  |  ||||| ||||||
Db    267 AACGTGGCATTCCAGTCGCCGCAGGACGACCTCACCGAGATCCGCCCAGAAGACGCCGAG 208

Qy    517 gac 519
            | |
Db    207 GCC 205
```

```
RESULT  29
AAF21735
ID   AAF21735 standard; DNA; 1139 BP.
XX
AC   AAF21735;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Human breast and ovarian cancer associated antigen gene SEQ ID 122.
XX
KW   Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW   nootropic; neurpprotective; antiviral; antiallergic; hepatotropic;
KW   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW   antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW   Addison's disease; allergy; autoimmune haemolytic anaemia;
KW   autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW   multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW   cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO200055173-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05881.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-611515/58.
DR   P-PSDB; AAB58832.
XX
PT   New human breast and ovarian cancer associated gene sequences and the
PT   polypeptides encoded by these genes, useful in the prevention,
PT   treatment and diagnosis of cancer, immune disorders, cardiovascular
PT   disorders and neurological diseases -
XX
PS   Claim 1; Page 573; 1299pp; English.
XX
CC   Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC   proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC   associated with breast and ovarian cancer. Included in the invention are
CC   sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC   isolation and characterisation of the DNA and protein sequences of the
CC   invention. The breast and ovarian cancer associated DNA, protein, agonist
CC   or antagonist sequences exhibit cytostatic; immunosuppressive;
CC   nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
CC   antibacterial; antifungal; antiparasitic and cardiant activity. The
CC   polynucleotide and protein sequences are used in the diagnosis of cancer,
CC   particularly breast and ovarian cancer. The nucleic acid sequences,
CC   proteins, agonists and agonists may also be used in the diagnosis,
CC   prevention and treatment of immune disorders e.g. Addison's disease,
CC   allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC   diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC   arthritis and ulcerative colitis; cardiovascular disorders such as
CC   myocardial ischaemias; wound healing; neurological diseases such as
CC   cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ   Sequence 1139 BP; 217 A; 350 C; 375 G; 188 T; 9 other;

  Query Match             2.6%;  Score 35;  DB 21;  Length 1139;
  Best Local Similarity   59.6%;  Pred. No. 3.5;
  Matches   59;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy   1242 ggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggccc 1301
           || ||  || ||    || |   || || |   |  |||||||| || |||  |  |||
```

```
Db    353 ggccgccgaggatggggcctgcggctgcctgcccaacccgggcacattcgaggagtgcca 412

Qy   1302 gctgttgcgcaaagtgctggaaaccattgagaaggaagg 1340
          | |  | |||| | ||||    ||||| ||| ||| ||
Db    413 ccggaagtgcaaggagctgtttcccattcagatggaggg 451


RESULT 30
AAS86098
ID   AAS86098 standard; cDNA; 1689 BP.
XX
AC   AAS86098;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #21902.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG21911.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 21902; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1689 BP; 348 A; 538 C; 525 G; 278 T; 0 other;


 Query Match             2.6%;  Score 35;  DB 23;  Length 1689;
 Best Local Similarity   59.6%;  Pred. No. 4.2;
 Matches   59;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy   1242 ggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggccc 1301
           || ||  || ||    || |   || || |   |  |||||||| || |||  |  |||
Db    285 ggccgccgaggatggggcctgcggctgcctgcccaacccgggcacattcgaggagtgcca 344

Qy   1302 gctgttgcgcaaagtgctggaaaccattgagaaggaagg 1340
          | |  | |||| | ||||    ||||| ||| ||| ||
Db    345 ccggaagtgcaaggagctgtttcccattcagatggaggg 383


RESULT 31
AAD08837
ID   AAD08837 standard; cDNA; 1325 BP.
XX
AC   AAD08837;
XX
DT   04-SEP-2001  (first entry)
XX
DE   Human G-protein coupled receptor-3 (GCREC-3) cDNA.
XX
KW   Human; G-protein coupled receptor-3; GCREC-3; gene therapy; cirrhosis;
KW   transgenic animal; proliferative disorder; actinic keratosis; hepatitis
KW   nephrotropic; cancer; breast; bladder; bone marrow; brain; uterus;
KW   leukaemia; adenocarcinoma; lymphoma; melanoma; myeloma; epilepsy; stroke;
KW   neurological disorder; Alzheimer's disease; Parkinson's disease; nausea;
KW   Huntington's disease; multiple sclerosis; dementia; angina pectoris;
KW   central nervous system disorder; cardiovascular disorder; hypertension;
KW   atherosclerosis; congestive heart failure; gastrointestinal disorder;
KW   dysphagia; peptic oesophagitis; spasm; gastritis; anorexia; pyrosis;
KW   pancreatitis; Crohn's disease; diarrhoea; autoimmune disorder; anaemia;
KW   inflammatory disorder; Acquired Immune Deficiency Syndrome; AIDS;
KW   Addison's diseases; allergy; asthma; diabetes mellitus; antithyroid;
KW   atopic dermatitis; glomerulonephritis; Grave's disease; psoriasis;
KW   rheumatoid arthritis; ulcerative colitis; osteoporosis; antifungal;
KW   metabolic disorder; obesity; nootropic; protozoacide; virucide; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             42..1310
FT                   /*tag=  a
FT                   /product= "Human GCREC-3 protein"
XX
PN   WO200142288-A2.
XX
PD   14-JUN-2001.
XX
PF   07-DEC-2000; 2000WO-US33382.
XX
PR   10-DEC-1999;   99US-0172852.
PR   22-DEC-1999;   99US-0171732.
PR   14-JAN-2000; 2000US-0176148.
PR   21-JAN-2000; 2000US-0177331.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Burford N,  Baughn MR,  Au-Young J,  Yang J,  Lu DAM,  Reddy R;
XX
DR   WPI; 2001-381635/40.
DR   P-PSDB; AAE04547.
XX
PT   New human G-protein coupled receptor polypeptides for diagnosing,
PT   preventing, and treating cell proliferative, neurological,
PT   cardiovascular, gastrointestinal, autoimmune and metabolic disorders -
XX
PS   Claim 5; Page 160-161; 175pp; English.
XX
CC   The present sequence is human G-protein coupled receptor-3 (GCREC-3)
CC   cDNA. GCREC is useful in somatic or germline gene therapy to correct a
CC   genetic deficiency, to express a conditionally lethal gene product and
CC   to express a protein which affords protection against intracellular
CC   parasites and also for diagnosis of disorders associated with expression
CC   of GCREC. GCREC is also useful for generating hybridisation probes useful
```

```
CC   in mapping the naturally occurring genomic sequences and to create
CC   knockin humanised animals (pigs) or transgenic animals (mice or rats) to
CC   model human diseases. GCREC is used to diagnose, prevent and treat
CC   proliferative disorders (actinic keratosis, arteriosclerosis, cirrhosis,
CC   hepatitis and cancer); cancer (breast, bladder, bone marrow, brain,
CC   uterus cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma)
CC   neurological disorders (epilepsy, stroke, Alzheimer's, Huntington's,
CC   Parkinson's disease, multiple sclerosis, dementia and other central
CC   nervous system disorders); cardiovascular disorders (angina pectoris,
CC   hypertension, atherosclerosis, congestive heart failure);
CC   gastrointestinal disorders (dysphagia, peptic oesophagitis, oesophageal
CC   spasm, gastritis, gastric carcinoma, anorexia, nausea, abdominal angina,
CC   pyrosis, pancreatitis, Crohn's disease, diarrhoea); autoimmune/
CC   inflammatory disorders (acquired immunodeficiency syndrome (AIDS),
CC   Addison's diseases, allergies, anaemia, asthma, diabetes mellitus, atopic
CC   dermatitis, glomerulonephritis, Grave's disease, osteoarthritis,
CC   psoriasis, rheumatoid arthritis, ulcerative colitis, bacterial, fungal,
CC   parasitic, protozoal and helminthic infections) and metabolic disorders
CC   (obesity, osteoporosis, viral infections).
XX
SQ   Sequence 1325 BP; 195 A; 473 C; 403 G; 254 T; 0 other;

  Query Match             2.6%;  Score 34.8;  DB 22;  Length 1325;
  Best Local Similarity   51.3%;  Pred. No. 4.3;
  Matches   81;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy     973 tccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgccgcagccggac 1032
           | | ||| ||  || ||| | ||||   ||| |||| ||   | ||||| | || |
Db     933 tactttagcgccgagggcttccgcaacaccctgcgcggcctgggcactccgcaccgggcc 992

Qy    1033 aaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaactcttcaagttc 1092
           |   |||  |||   |  || | |     ||    ||  | ||   | |    ||
Db     993 aggacctcggccaccaacgggacgcgggcggcgctcgcgcaatccgaaaggtccgccgtc 1052

Qy    1093 acgacagccgtcaacggcggcgaccgcgccatggtacc 1130
           || || | || || | |    ||   |||||||  ||
Db    1053 accaccgacgccaccaggccggatgccgccatgtcccc 1090


RESULT  32
AAS85732/c
ID   AAS85732 standard; cDNA; 1743 BP.
XX
AC   AAS85732;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #21536.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG21545.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 21536; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1743 BP; 384 A; 543 C; 442 G; 374 T; 0 other;

  Query Match             2.6%;  Score 34.8;  DB 23;  Length 1743;
  Best Local Similarity   52.0%;  Pred. No. 4.9;
  Matches   78;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     962 acaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgc 1021
           | || | || |  |   |||   || | | || |||  ||   |||  ||  | ||||||
Db    1637 ATAAGCTGAATATCTGGATCCTTGATGACGGCGGCAGGGAAGAGTTTCGCCAGTTTGCGC 1578

Qy    1022 cgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaac 1081
           |   ||      | |  ||| | ||||||||||  |  |||     |  |  |||
Db    1577 AAAACGTGGGGGTGAAATATATCGCCCGCACCACTCATGAACATGCGAAAGCAGGCAACA 1518

Qy    1082 tcttcaagttcacgacagccgtcaacggcg 1111
           ||  |||      || |   | ||| ||||
Db    1517 TCAACAATGCGCTGAAATATGCCAAAGGCG 1488


RESULT  33
AAS85741/c
ID   AAS85741 standard; cDNA; 1875 BP.
XX
AC   AAS85741;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #21545.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
```

```
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG21554.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 21545; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1875 BP; 393 A; 554 C; 535 G; 393 T; 0 other;


  Query Match             2.6%;  Score 34.8;  DB 23;  Length 1875;
  Best Local Similarity   52.0%;  Pred. No. 5.1;
  Matches   78;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     962 acaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgc 1021
           | || | || |  |   |||   || | | || ||| ||   |||  ||  | ||||||
Db     800 ATAAGCTGAATATCTGGATCCTTGATGACGGCGGCAGGGAAGAGTTTCGCCAGTTTGCGC 741

Qy    1022 cgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaac 1081
             |   ||      | |  ||| | ||||||||||  |  |||     |  |  |||
Db     740 AAAACGTGGGGGTGAAATATATCGCCCGCACCACTCATGAACATGCGAAAGCAGGCAACA 681

Qy    1082 tcttcaagttcacgacagccgtcaacggcg 1111
           ||  |||      || |   | ||| ||||
Db     680 TCAACAATGCGCTGAAATATGCCAAAGGCG 651


RESULT  34
AAS82534
ID   AAS82534 standard; cDNA; 2196 BP.
XX
AC   AAS82534;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #18338.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG18347.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 18338; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2196 BP; 497 A; 570 C; 620 G; 509 T; 0 other;


  Query Match             2.6%;  Score 34.8;  DB 23;  Length 2196;
  Best Local Similarity   52.0%;  Pred. No. 5.5;
  Matches   78;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     962 acaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgc 1021
           | || | || |  |   |||   || | | || ||| ||   |||  ||  | ||||||
Db     914 ataagctgaatatctggatccttgatgacggcggcagggaagagtttcgccagtttgcgc 973

Qy    1022 cgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaac 1081
             |   ||      | |  ||| | ||||||||||  |  |||     |  |  |||
Db     974 aaaacgtgggggtgaaatatatcgcccgcaccactcatgaacatgcgaaagcaggcaaca 1033

Qy    1082 tcttcaagttcacgacagccgtcaacggcg 1111
           ||  |||      || |   | ||| ||||
Db    1034 tcaacaatgcgctgaaatatgccaaaggcg 1063


RESULT  35
AAS85733/c
ID   AAS85733 standard; cDNA; 2196 BP.
XX
AC   AAS85733;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #21537.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
```

XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG21546.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 21537; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2196 BP; 509 A; 620 C; 570 G; 497 T; 0 other;

```
  Query Match             2.6%;  Score 34.8;  DB 23;  Length 2196;
  Best Local Similarity   52.0%;  Pred. No. 5.5;
  Matches   78;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     962 acaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgc 1021
           | || | || |  |   |||   || | | || |||  ||   |||  ||  | ||||||
Db    1283 ATAAGCTGAATATCTGGATCCTTGATGACGGCGGCAGGGAAGAGTTTCGCCAGTTTGCGC 1224

Qy    1022 cgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaac 1081
             |   ||      | |  ||| | ||||||||||  |  |||     |  |  |||
Db    1223 AAAACGTGGGGGTGAAATATATCGCCCGCACCACTCATGAACATGCGAAAGCAGGCAACA 1164

Qy    1082 tcttcaagttcacgacagccgtcaacggcg 1111
           ||  |||      || |   | ||| ||||
Db    1163 TCAACAATGCGCTGAAATATGCCAAAGGCG 1134
```

RESULT 36
AAS89351
ID AAS89351 standard; cDNA; 2196 BP.
XX
AC AAS89351;

XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #25155.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG25164.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 25155; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2196 BP; 497 A; 570 C; 620 G; 509 T; 0 other;

```
  Query Match             2.6%;  Score 34.8;  DB 23;  Length 2196;
  Best Local Similarity   52.0%;  Pred. No. 5.5;
  Matches   78;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     962 acaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgc 1021
           | || | || |  |   |||   || | | || |||  ||   |||  ||  | ||||||
Db     914 ataagctgaatatctggatccttgatgacggcggcagggaagagtttcgccagtttgcgc 973

Qy    1022 cgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaac 1081
             |   ||      | |  ||| | ||||||||||  |  |||     |  |  |||
Db     974 aaaacgtgggggtgaaatatatcgcccgcaccactcatgaacatgcgaaagcaggcaaca 1033

Qy    1082 tcttcaagttcacgacagccgtcaacggcg 1111
           ||  |||      || |   | ||| ||||
Db    1034 tcaacaatgcgctgaaatatgccaaaggcg 1063
```

```
RESULT  37
AAS73293
ID   AAS73293 standard; cDNA; 2909 BP.
XX
AC   AAS73293;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #9097.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG09106.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 9097; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2909 BP; 687 A; 728 C; 811 G; 683 T; 0 other;


  Query Match             2.6%;  Score 34.8;  DB 23;  Length 2909;
  Best Local Similarity   52.0%;  Pred. No. 6.3;
  Matches   78;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     962 acaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgc 1021
           | || | || |  |   |||   || | | || |||  ||   |||  ||  | ||||||
Db     914 ataagctgaatatctggatccttgatgacggcggcagggaagagtttcgccagtttgcgc 973

Qy    1022 cgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaac 1081
             |  ||      | |  ||| | ||||||||||  |  |||     |  |  |||
Db     974 aaaacgtgggggtgaaatatatcgcccgcaccactcatgaacatgcgaaagcaggcaaca 1033

Qy    1082 tcttcaagttcacgacagccgtcaacggcg 1111
           ||  |||      || |   | ||| ||||
Db    1034 tcaacaatgcgctgaaatatgccaaaggcg 1063


RESULT  38
AAS82533
ID   AAS82533 standard; cDNA; 3810 BP.
XX
AC   AAS82533;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #18337.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG18346.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 18337; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3810 BP; 894 A; 1035 C; 1032 G; 849 T; 0 other;


  Query Match             2.6%;  Score 34.8;  DB 23;  Length 3810;
```

```
  Best Local Similarity   52.0%;  Pred. No. 7.2;
  Matches   78;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     962 acaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgc 1021
           | || | || |  |   |||   || | | || ||| ||  |||  ||  | ||||||
Db    1244 ataagctgaatatctggatccttgatgacggcggcagggaagagtttcgccagtttgcgc 1303

Qy    1022 cgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaac 1081
             |  ||      | |  ||| | ||||||||||  |  |||     |  |  |||
Db    1304 aaaacgtgggggtgaaatatatcgcccgcaccactcatgaacatgcgaaagcaggcaaca 1363

Qy    1082 tcttcaagttcacgacagccgtcaacggcg 1111
           ||  |||      || |   | ||| ||||
Db    1364 tcaacaatgcgctgaaatatgccaaaggcg 1393
```

```
RESULT  39
AAS89350
ID   AAS89350 standard; cDNA; 3810 BP.
XX
AC   AAS89350;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #25154.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG25163.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 25154; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3810 BP; 894 A; 1035 C; 1032 G; 849 T; 0 other;
```

```
  Query Match             2.6%;  Score 34.8;  DB 23;  Length 3810;
  Best Local Similarity   52.0%;  Pred. No. 7.2;
  Matches   78;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     962 acaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgc 1021
           | || | || |  |   |||   || | | || ||| ||  |||  ||  | ||||||
Db    1244 ataagctgaatatctggatccttgatgacggcggcagggaagagtttcgccagtttgcgc 1303

Qy    1022 cgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaac 1081
             |  ||      | |  ||| | ||||||||||  |  |||     |  |  |||
Db    1304 aaaacgtgggggtgaaatatatcgcccgcaccactcatgaacatgcgaaagcaggcaaca 1363

Qy    1082 tcttcaagttcacgacagccgtcaacggcg 1111
           ||  |||      || |   | ||| ||||
Db    1364 tcaacaatgcgctgaaatatgccaaaggcg 1393
```

```
RESULT  40
ABL02343/c
ID   ABL02343 standard; cDNA; 783 BP.
XX
AC   ABL02343;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 1511.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB58240.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 1511; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 783 BP; 173 A; 201 C; 229 G; 180 T; 0 other;
```

```
Query Match             2.6%;  Score 34.4;  DB 23;  Length 783;
Best Local Similarity   52.9%;  Pred. No. 4.4;
Matches   74;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy     849 ggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtgatttccggttc 908
           ||||| | | | | || | | || |   | |||||   ||| ||||  || |  | |
Db     499 GGTGTTTGATCCTTCCCCAGCCGGAAATGATGACGTCCACATCCGCTGGAGTATCCGCAG 440

Qy     909 ggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgctaccacaatca 968
             |  | || || || | |       |||  || ||| | | ||  | |||     | ||
Db     439 TGGGTAGATCGATGGGCTGAATACTCGCGGACAGGATCAGTGGGCTCTCTAGTCGCAGCA 380

Qy     969 gatttccgttatcgaagaag 988
           ||    |||  |  | ||||
Db     379 GAGCAACGTCGTTAAGGAAG 360


RESULT  41
ABL02342/c
ID   ABL02342 standard; cDNA; 2783 BP.
XX
AC   ABL02342;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 1508.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB58239.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 1508; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2783 BP; 842 A; 566 C; 549 G; 826 T; 0 other;


Query Match             2.6%;  Score 34.4;  DB 23;  Length 2783;
Best Local Similarity   52.9%;  Pred. No. 8.2;
Matches   74;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy     849 ggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtgatttccggttc 908
           ||||| | | | | || | | || |   | |||||   ||| ||||  || |  | |
Db    1499 GGTGTTTGATCCTTCCCCAGCCGGAAATGATGACGTCCACATCCGCTGGAGTATCCGCAG 1440

Qy     909 ggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgctaccacaatca 968
             |  | || || || | |       |||  || ||| | | ||  | |||     | ||
Db    1439 TGGGTAGATCGATGGGCTGAATACTCGCGGACAGGATCAGTGGGCTCTCTAGTCGCAGCA 1380

Qy     969 gatttccgttatcgaagaag 988
           ||    |||  |  | ||||
Db    1379 GAGCAACGTCGTTAAGGAAG 1360


RESULT  42
AAS59558/c
ID   AAS59558 standard; DNA; 25365 BP.
XX
AC   AAS59558;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Propionibacterium acnes immunogenic protein encoding DNA #53.
XX
KW   SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW   dermatological; osteopathic; neuroprotectant; ds.
XX
OS   Propionibacterium acnes.
XX
PN   WO200181581-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US12865.
XX
PR   21-APR-2000; 2000US-199047P.
PR   02-JUN-2000; 2000US-208841P.
PR   07-JUL-2000; 2000US-216747P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Persing DH,  Mitcham JL,  Wang SS,  Bhatia A;
PI   L'maisonneuve J,  Zhang Y,  Jen S,  Carter D;
XX
DR   WPI; 2001-616774/71.
XX
PT   Propionibacterium acnes polypeptides and nucleic acids useful for
PT   vaccinating against and diagnosing infections, especially useful for
PT   treating acne vulgaris -
XX
PS   Claim 1; SEQ ID No 53; 1069pp; English.
XX
CC   Sequences AAS59506-AAS59804 represent DNA molecules encoding
CC   Propionibacterium acnes immunogenic polypeptides. The proteins and their
CC   associated DNA sequences are used in the treatment, prevention and
CC   diagnosis of medical conditions caused by P. acnes. The disorders include
CC   SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
CC   osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved
CC   in infections of bone, joints and the central nervous system, however it
CC   is particularly involved in the inflammatory lesions associated with acne
CC   vulgaris. A method for detecting the presence or absence of P. acnes in a
CC   patient comprises contacting a sample with a binding agent that binds to
CC   the proteins of the invention and determining the amount of bound protein
CC   in the sample. The polypeptides may be used as antigens in the production
CC   of antibodies specific for P. acnes proteins. These antibodies can be
CC   used to downregulate expression and activity of P. acnes polypeptides and
CC   therefore treat P. acnes infections. The antibodies may also be used as
CC   diagnostic agents for determining P. acnes presence, for example, by
CC   enzyme linked immunosorbent assay (ELISA). This sequence encodes the
CC   polypeptides shown in AAU52925-AAU53195 and AAU67543-AAU67545.
```

```
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 25365 BP; 5364 A; 7871 C; 7324 G; 4798 T; 8 other;


  Query Match             2.5%;  Score 34.2;  DB 23;  Length 25365;
  Best Local Similarity   55.5%;  Pred. No. 28;
  Matches   66;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy    1164 catcctgcctaccttgcttttgcgcgatttaatcgtcggcgataccgacagcgcgcaggc 1223
           | || || | ||   | | |||||| || | |||  | || || || | |    | |
Db   16596 CGTCATGACGACTGAGGTATTGCGCAATATGATCTACCGCAATTCCCATACTCTCGACAC 16537

Qy    1224 tttgggttgcttggaattggacgaagaagacctcgctttgtgcagcttcgtctgcccgg 1282
            ||||| ||  |||   |||| || |   ||  |     |  | | |||  | ||||||
Db   16536 CTTGGGGTGGGTGGTGCTGGATGAGGTGCACTACCTCGCGGACCGTTTCCGCGGCCCGG 16478


RESULT  43
AAS54057
ID   AAS54057 standard; DNA; 1191 BP.
XX
AC   AAS54057;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Pseudomonas aeruginosa DNA for cellular proliferation protein #188.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Pseudomonas aeruginosa.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU36198.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 7694; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1191 BP; 222 A; 420 C; 365 G; 184 T; 0 other;


  Query Match             2.5%;  Score 34;  DB 23;  Length 1191;
  Best Local Similarity   52.1%;  Pred. No. 7.2;
  Matches   76;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy     427 gatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggctgccgac 486
           || | || |  |      ||| ||    | | | ||    || | | ||||      |||
Db     604 gacggcgtgatggagctgatcatcaagcaggtgctgccggccgagctgctgcacaaggac 663

Qy     487 cctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggtattgagc 546
             |   | ||  ||||||     ||| | |   | | |   ||||| | | |      ||
Db     664 acccagttccacatcaacccgaccggcaacttcgtgatcggcggcccggtcggcgactgc 723

Qy     547 cgcctgaccgaacgtaaaatccatgt 572
           |||||||||| ||| || |||   ||
Db     724 ggcctgaccggacgcaagatcatcgt 749


RESULT  44
AAS73398/c
ID   AAS73398 standard; cDNA; 2367 BP.
XX
AC   AAS73398;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #9202.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG09211.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 9202; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
```

```
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2367 BP; 530 A; 691 C; 645 G; 501 T; 0 other;


  Query Match             2.5%;  Score 33.8;  DB 23;  Length 2367;
  Best Local Similarity   54.4%;  Pred. No. 12;
  Matches   68;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy     186 gaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcaccgtgg 245
           | | |  |||||| |  |   |||| |||     ||||||||||| |       || |
Db     859 GGAACACGGCGTATTTATCGGTGCGGCGGGCGATGGCAAAATCGCCTCAGCAACGCGGGC 800

Qy     246 cgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatcgagtt 305
            ||  |    | |  |  | ||  |||||||  ||  |  ||  |||| |  |  |  |
Db     799 AGATTATGCGGCAGCTGCGGCACGCGTGATTAGCGAAGCCGGTCACGAAGGCAAGGTTTA 740

Qy     306 cgaac 310
           |||||
Db     739 CGAAC 735


RESULT  45
AAS73427/c
ID   AAS73427 standard; cDNA; 2367 BP.
XX
AC   AAS73427;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #9231.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG09240.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 9231; 103pp; English.

XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2367 BP; 530 A; 691 C; 645 G; 501 T; 0 other;


  Query Match             2.5%;  Score 33.8;  DB 23;  Length 2367;
  Best Local Similarity   54.4%;  Pred. No. 12;
  Matches   68;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy     186 gaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcaccgtgg 245
           | | |  |||||| |  |   |||| |||     ||||||||||| |       || |
Db     859 GGAACACGGCGTATTTATCGGTGCGGCGGGCGATGGCAAAATCGCCTCAGCAACGCGGGC 800

Qy     246 cgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatcgagtt 305
            ||  |    | |  |  | ||  |||||||  ||  |  ||  |||| |  |  |  |
Db     799 AGATTATGCGGCAGCTGCGGCACGCGTGATTAGCGAAGCCGGTCACGAAGGCAAGGTTTA 740

Qy     306 cgaac 310
           |||||
Db     739 CGAAC 735


Search completed: June 30, 2002, 13:46:02
Job time: 67395 sec
```